

## PROBABILITY BOXES ON TOTALLY PREORDERED SPACES FOR MULTIVARIATE MODELLING

MATTHIAS TROFFAES AND SEBASTIEN DESTERCCKE

**ABSTRACT.** A pair of lower and upper cumulative distribution functions, also called probability box or p-box, is among the most popular models used in imprecise probability theory. They arise naturally in expert elicitation, for instance in cases where bounds are specified on the quantiles of a random variable, or when quantiles are specified only at a finite number of points. Many practical and formal results concerning p-boxes already exist in the literature. In this paper, we provide new efficient tools to construct multivariate p-boxes and develop algorithms to draw inferences from them. For this purpose, we formalise and extend the theory of p-boxes using Walley's behavioural theory of imprecise probabilities, and heavily rely on its notion of natural extension and existing results about independence modeling. In particular, we allow p-boxes to be defined on arbitrary totally preordered spaces, hence thereby also admitting multivariate p-boxes via probability bounds over any collection of nested sets. We focus on the cases of independence (using the factorization property), and of unknown dependence (using the Fréchet bounds), and we show that our approach extends the probabilistic arithmetic of Williamson and Downs. Two design problems—a damped oscillator, and a river dike—demonstrate the practical feasibility of our results.

### 1. INTRODUCTION

Imprecise probability [41] refers to uncertainty models applicable in situations where the available information does not allow us to single out a unique probability measure for all random variables involved. Examples of such models include 2- and  $n$ -monotone capacities [5], lower and upper previsions [43, 44, 41], belief functions [38], credal sets [31], possibility and necessity measures [23, 6], interval probabilities [42], and coherent risk measures [1, 15].

Unlike classical probability models, which are described by probability measures, imprecise probability models require more complex mathematical tools, such as non-linear functionals and non-additive measures [41]. It is therefore of interest to consider particular imprecise probability models that yield simple mathematical descriptions, possibly at the expense of generality, but gaining ease of use, elicitation, and graphical representation.

One such model is considered in this paper: pairs of lower and upper distribution functions, also called *probability boxes*, or briefly, p-boxes [24, 25]. P-boxes are often used in risk or safety studies, in which cumulative distributions play an essential role. Many theoretical properties and practical aspects of p-boxes have already been studied in the literature. Previous work includes probabilistic arithmetic [45], which provides a very efficient numerical framework for particular statistical inferences with p-boxes (and which we generalise in this paper). In [26], p-boxes are connected to info-gap theory [2]. The relation between p-boxes and random sets was investigated in [30] and applied in [35]: many results and techniques applicable to random sets are also applicable to p-boxes. Finally, a recent extension of p-boxes to arbitrary finite spaces [19] yields potential application of p-boxes to a much more general set of problems, such as robust design analysis [20, 28], and signal processing [21].

---

*Key words and phrases.* Lower prevision, p-box, multivariate, Choquet integral, Fréchet bounds, full component.

In this paper, we study p-boxes within the framework of the theory of coherent lower previsions. Coherent lower previsions were introduced by Williams [43] as a generalisation of de Finetti's work [11], and were developed further by Walley [41]. Coherent lower previsions generalize many of the other imprecise probability models in the literature, and are equivalent to closed convex sets of finitely additive probability measures.

Studying p-boxes by means of lower previsions has at least two advantages:

- Lower previsions can be defined on arbitrary spaces, and thus enable p-boxes to be used for more general problems, and not just problems concerning the real line.
- Lower previsions come with a powerful inference tool, called *natural extension*, which reflects the least-committal consequences of any given assessments. Natural extension generalises many known extensions, including for instance Choquet integration for 2-monotone measures. In this paper, we study the natural extension of a p-box, and we derive a number of useful expressions for it. This leads to new numerical tools that provide exact inferences on arbitrary events, and even on arbitrary (bounded) random quantities.

From the point of view of coherent lower previsions, p-boxes have already been studied briefly in [41, Section 4.6.6] and [39]. Lower and upper distribution functions associated with a sequence of moments have also been considered [33].

As already mentioned, [19] extended p-boxes to finite totally preordered spaces. In this paper, we extend p-boxes further to arbitrary totally preordered spaces. Our generalisation has many useful features that classical p-boxes do not have:

- We encompass, in one sweep, p-boxes defined on finite spaces, as well as (continuous) p-boxes on closed real intervals.
- Perhaps even more importantly, as we do not impose anti-symmetry on the ordering, we can also handle product spaces by considering an appropriate total preorder—for instance, one induced by a metric—and thus also admit multivariate non-finite p-boxes, which have not been considered before. Whence, we can specify p-boxes directly on the product space. Contrast this with the usual multivariate approach to p-boxes, such as probabilistic arithmetic [45], that consider one marginal p-box per dimension and draw inferences from a joint model built around some information about variable dependencies (of course, we can still do the same, and will do so in Section 7).
- Our approach is also useful in elicitation, as it allows uncertainty to be expressed as probability bounds over any collection of (possibly multivariate) nested sets—also see [28, 18] for a discussion of similarly constructed models. Thus, unlike classical p-boxes, we are not restricted to events of the type  $[-\infty, x]$ , even on the real line.

Our approach is thus rather different, and far more general, than the one usually considered for inferences with p-boxes. Indeed, we first define a joint p-box over some multivariate space of interest, either directly or by using marginal models and a dependence model, after which we draw exact inferences from this joint p-box using natural extension. In contrast, usual methods [36] such as for instance probabilistic arithmetic [45] start out with marginal p-boxes each defined on the real line, and provide tools to make inferences for specific multivariate events.

The paper is organised as follows: Section 2 provides a brief introduction to the theory of coherent lower previsions, used in the rest of the paper. Section 3 then introduces and studies the p-box model from the point of view of lower previsions. Section 4 provides an expression for the natural extension of a p-box to all events, via the partition topology induced by the equivalence classes of the preorder and additivity on full components. Section 5 studies the natural extension to all gambles, via the Choquet integral. Section 6 studies an important special case of p-boxes whose preorder is induced by a real-valued mapping, as this will usually be the

most convenient way to specify a multivariate p-box. Section 7 discusses the construction of such multivariate p-boxes from marginal coherent lower previsions under arbitrary dependency models. For two important special cases—epistemic independence, and completely unknown dependence (the Fréchet case)—closed expressions are derived. We also derive probabilistic arithmetic as a special case of our approach. Section 8 demonstrates the theory with some examples. As a first example, we infer bounds on the expectation of the damping ratio of a damped harmonic oscillator whose parameters are described by a multivariate p-box. As a second example, we derive the expected overflow height for a river dike, again using a multivariate p-box. Finally, Section 9 ends with our main conclusions and open problems.

## 2. PRELIMINARIES

We start with a brief introduction to the imprecise probability models that we shall use in this paper. We refer to [4], [43] and [41] for more details. See also [32] for a brief summary of the theory.

Let  $\Omega$  be the possibility space. A subset of  $\Omega$  is called an *event*. A *gamble* on  $\Omega$  is a bounded real-valued function on  $\Omega$ . The set of all gambles on  $\Omega$  is denoted by  $\mathcal{L}(\Omega)$ , or simply by  $\mathcal{L}$  if the possibility space is clear from the context. A particular type of gamble is the *indicator* of an event  $A$ , which is the gamble that takes the value 1 on elements of  $A$  and the value 0 elsewhere, and is denoted by  $I_A$ , or simply by  $A$  if no confusion with  $A$  as event is possible.

A *lower prevision*  $\underline{P}$  is a real-valued functional defined on an arbitrary subset  $\mathcal{K}$  of  $\mathcal{L}$ , and is interpreted as follows: for any gamble  $f$  in  $\mathcal{K}$  and any  $\epsilon > 0$ , the transaction  $f - \underline{P}(f) + \epsilon$  is acceptable to the subject who has assessed this lower prevision. Hence, lower previsions summarize a subject's supremum buying prices for a collection of gambles, and it can be argued that in this way they model a subject's belief about the true state  $x$  in  $\Omega$  (see [41] for a more in-depth explanation). A lower prevision defined on a set of indicators of events is usually called a *lower probability*.

By  $\overline{P}$ , we denote the conjugate *upper prevision* of  $\underline{P}$ : for every  $f$  such that  $-f \in \mathcal{K}$ ,  $\overline{P}(f) = -\underline{P}(-f)$ . The upper prevision  $\overline{P}(f)$  can be interpreted as a subject's infimum selling price for  $f$ .

A lower prevision on  $\mathcal{K}$  is called *coherent* (see [43, p. 5] and [41, pp. 73–75, Sec. 2.5]) when for all  $p$  in  $\mathbb{N}$ , all  $f_0, f_1, \dots, f_p$  in  $\mathcal{K}$  and all non-negative real numbers  $\lambda_0, \lambda_1, \dots, \lambda_p$ ,

$$\sup_{x \in \Omega} \left[ \sum_{i=1}^p \lambda_i (f_i(x) - \underline{P}(f_i)) - \lambda_0 (f_0(x) - \underline{P}(f_0)) \right] \geq 0.$$

A lower prevision on the set  $\mathcal{L}$  of all gambles is coherent if and only if (see [43, p. 11, Sec. 1.2.2] and [41, p. 75, Sec. 2.5.5])

- (C1)  $\underline{P}(f) \geq \inf f$ ,
- (C2)  $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ , and
- (C3)  $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$

for all gambles  $f, g$  and all non-negative real numbers  $\lambda$ .

A functional on  $\mathcal{L}$  satisfying  $P(f) \geq \inf f$  and  $P(f + g) = P(f) + P(g)$  for any pair of gambles  $f$  and  $g$  is called a *linear prevision* on  $\mathcal{L}$  [41, p. 88, Sec. 2.4.8], and the set of all linear previsions on  $\mathcal{L}$  is denoted by  $\mathcal{P}$ . A linear prevision is the expectation operator with respect to its restriction to events, which is a finitely additive probability.

An alternative characterisation of coherence via linear previsions goes as follows. Let  $\underline{P}$  be a lower prevision on  $\mathcal{K}$  and let  $\mathcal{M}(\underline{P})$  denote the set of all linear previsions on  $\mathcal{L}$  that dominate  $\underline{P}$  on  $\mathcal{K}$ :

$$\mathcal{M}(\underline{P}) = \{Q \in \mathcal{P} : (\forall f \in \mathcal{K})(Q(f) \geq \underline{P}(f))\}$$

Then  $\underline{P}$  is coherent if and only if  $\underline{P}$  agrees with the lower envelope of  $\mathcal{M}(\underline{P})$  on  $\mathcal{K}$ , that is, if and only if  $\underline{P}(f) = \min_{Q \in \mathcal{M}(\underline{P})} Q(f)$  for all  $f \in \mathcal{K}$  (see [43, p. 18] and [41, p. 138, Sec. 3.3.3]). A consequence of this is that a lower envelope of coherent lower previsions is again a coherent lower prevision.

Given a coherent lower prevision  $\underline{P}$  on  $\mathcal{K}$ , its *natural extension* [41, Chapter 3] to a larger set  $\mathcal{K}_1$  of gambles ( $\mathcal{K}_1 \supseteq \mathcal{K}$ ) is the pointwise smallest coherent lower prevision on  $\mathcal{K}_1$  that agrees with  $\underline{P}$  on  $\mathcal{K}$ . Because it is the pointwise smallest, the natural extension is the most conservative (or least-committal) coherent extension, and thereby reflects the minimal behavioural consequences of  $\underline{P}$  on  $\mathcal{K}_1$ .

Taking natural extension is transitive [39, p. 98, Cor. 4.9]: if  $\underline{E}_1$  is the natural extension of  $\underline{P}$  to  $\mathcal{K}_1$  and  $\underline{E}_2$  is the natural extension of  $\underline{E}_1$  to  $\mathcal{K}_2 \supseteq \mathcal{K}_1$ , then  $\underline{E}_2$  is also the natural extension of  $\underline{P}$  to  $\mathcal{K}_2$ . Hence, if we know the natural extension of  $\underline{P}$  to the set of all gambles then we also know the natural extension of  $\underline{P}$  to any set of gambles that includes  $\mathcal{K}$ . The natural extension to all gambles is usually denoted by  $\underline{E}$ . It holds that  $\underline{E}(f) = \min_{Q \in \mathcal{M}(\underline{P})} Q(f)$  for any  $f \in \mathcal{L}$  [41, p. 136, Sec. 3.4.1].

A particular class of coherent lower previsions of interest in this paper are *completely monotone lower previsions* [10, 9]. A lower prevision  $\underline{P}$  defined on a lattice of gambles  $\mathcal{K}$ , i.e., a set of gambles closed under point-wise maximum and point-wise minimum, is called  $n$ -monotone if for all  $p \in \mathbb{N}$ ,  $p \leq n$ , and all  $f, f_1, \dots, f_p$  in  $\mathcal{K}$ :

$$\sum_{I \subseteq \{1, \dots, p\}} (-1)^{|I|} \underline{P} \left( f \wedge \bigwedge_{i \in I} f_i \right) \geq 0.$$

A lower prevision which is  $n$ -monotone for all  $n \in \mathbb{N}$  is called *completely monotone*.

### 3. P-BOXES

In this section, we introduce the formalism of p-boxes defined on totally preordered (not necessarily finite) spaces, as well as the field of events  $\mathcal{H}$ , which will be instrumental to study the natural extension of p-boxes. Hence, in contrast to the work by [24], we do not restrict p-boxes to intervals on the real line.

Let  $(\Omega, \leq)$  be a total preorder: so  $\leq$  is transitive and reflexive and any two elements are comparable. We write  $x < y$  for  $x \leq y$  and  $x \not\leq y$ ,  $x > y$  for  $y < x$ , and  $x \simeq y$  for  $x \leq y$  and  $y \leq x$ . For any two  $x, y \in \Omega$  exactly one of  $x < y$ ,  $x \simeq y$ , or  $x > y$  holds. We also use the following common notation for intervals in  $\Omega$ :

$$\begin{aligned} [x, y] &= \{z \in \Omega : x \leq z \leq y\} \\ (x, y) &= \{z \in \Omega : x < z < y\} \end{aligned}$$

and similarly for  $[x, y)$  and  $(x, y]$ .

For simplicity, we assume that  $\Omega$  has a smallest element  $0_\Omega$  and a largest element  $1_\Omega$ . This is no essential assumption, since we can always add these two elements to the space  $\Omega$ .

A *cumulative distribution function* is a mapping  $F : \Omega \rightarrow [0, 1]$  which is non-decreasing and satisfies moreover  $F(1_\Omega) = 1$ . For each  $x \in \Omega$ , we interpret  $F(x)$  as the probability of the interval  $[0_\Omega, x]$ . Note that we do not impose  $F(0_\Omega) = 0$ , so we allow  $\{0_\Omega\}$  to carry non-zero mass, which happens commonly if  $\Omega$  is finite. Also note that distribution functions are not assumed to be right-continuous—this would make no sense since we have no topology defined yet on  $\Omega$ —but even if there is a topology for  $\Omega$ , we make no continuity assumptions.

By  $\Omega/\simeq$  we denote the quotient set of  $\Omega$  with respect to the equivalence relation  $\simeq$  induced by  $\leq$ , that is:

$$\begin{aligned} [x]_{\simeq} &= \{y \in \Omega : y \simeq x\} \text{ for any } x \in \Omega \\ \Omega/\simeq &= \{[x]_{\simeq} : x \in \Omega\} \end{aligned}$$

Because  $F$  is non-decreasing,  $F$  is constant on elements  $[x]_{\simeq}$  of  $\Omega/\simeq$ .

**Definition 1.** A *probability box*, or *p-box*, is a pair  $(\underline{F}, \overline{F})$  of cumulative distribution functions from  $\Omega$  to  $[0, 1]$  satisfying  $\underline{F} \leq \overline{F}$ .

Note that some definitions, such as in [19], differ in that they use the coarsest preorder for which  $\underline{F}$  and  $\overline{F}$  are non-decreasing, effectively imposing that  $x \simeq y$  if and only if  $\underline{F}(x) = \underline{F}(y)$  and  $\overline{F}(x) = \overline{F}(y)$ . This paper follows [17], only imposing that  $\underline{F}$  and  $\overline{F}$  are non-decreasing, so  $x \simeq y$  implies  $\underline{F}(x) = \underline{F}(y)$  and  $\overline{F}(x) = \overline{F}(y)$ , but not the other way around, thereby admitting more preorders, possibly leading to tighter bounds (see Example 6 further on).

A p-box is interpreted as a lower and an upper cumulative distribution function. In Walley's framework, this means that a p-box is interpreted as a lower probability  $\underline{P}_{\underline{F}, \overline{F}}$  on the set of events

$$\mathcal{K} = \{[0_{\Omega}, x] : x \in \Omega\} \cup \{(y, 1_{\Omega}] : y \in \Omega\}$$

by

$$\underline{P}_{\underline{F}, \overline{F}}([0_{\Omega}, x]) = \underline{F}(x) \text{ and } \underline{P}_{\underline{F}, \overline{F}}((y, 1_{\Omega}]) = 1 - \overline{F}(y).$$

In the particular case of p-boxes on  $[a, b] \subseteq \mathbb{R}$  it was mentioned by [41, Section 4.6.6] and proven by [39, p. 93, Thm. 3.59] that  $\underline{P}_{\underline{F}, \overline{F}}$  is coherent. More generally, it is straightforward to show that p-boxes on an arbitrary totally preordered space  $(\Omega, \leq)$  are coherent as well.<sup>1</sup> We denote by  $\underline{E}_{\underline{F}, \overline{F}}$  the natural extension of  $\underline{P}_{\underline{F}, \overline{F}}$  to all gambles. We study this natural extension extensively further on.

When  $\underline{F} = \overline{F}$ , we say that  $(\underline{F}, \overline{F})$  is *precise*, and we denote the corresponding lower prevision on  $\mathcal{K}$  by  $\underline{P}_F$  and its natural extension to  $\mathcal{L}$  by  $\underline{E}_F$  (with  $F := \underline{F} = \overline{F}$ ).

A few examples of p-boxes on  $[0, 1]$  are illustrated in Fig 1.

Given a p-box, we can consider the set of distribution functions that lie between  $\underline{F}$  and  $\overline{F}$ ,

$$\Phi(\underline{F}, \overline{F}) = \{F : \underline{F} \leq F \leq \overline{F}\}.$$

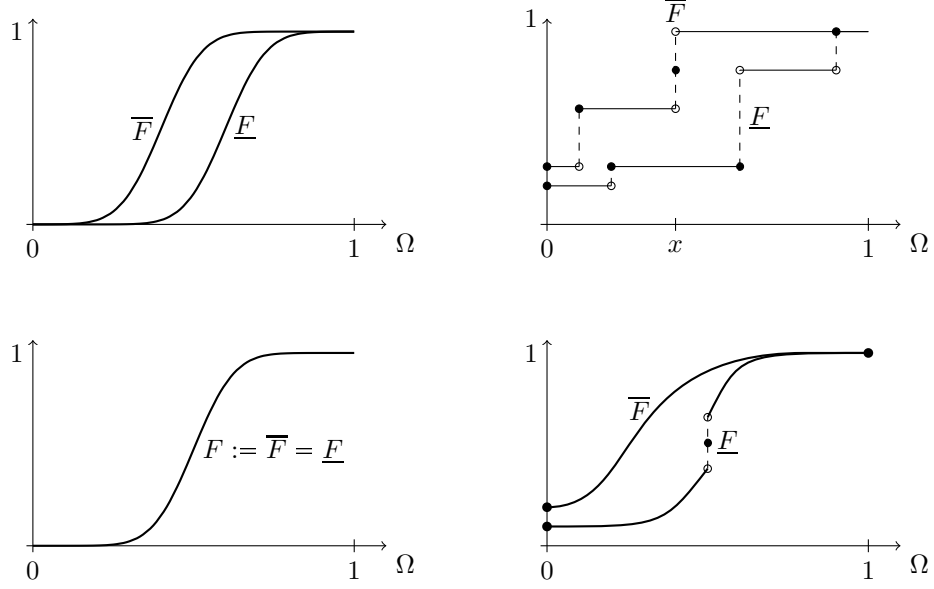
We can easily express the natural extension  $\underline{E}_{\underline{F}, \overline{F}}$  of  $\underline{P}_{\underline{F}, \overline{F}}$  in terms of  $\Phi(\underline{F}, \overline{F})$ :  $\underline{E}_{\underline{F}, \overline{F}}$  is the lower envelope of the natural extensions of the cumulative distribution functions<sup>2</sup>  $F$  that lie between  $\underline{F}$  and  $\overline{F}$ :

$$\underline{E}_{\underline{F}, \overline{F}}(f) = \inf_{F \in \Phi(\underline{F}, \overline{F})} \underline{E}_F(f) \quad (1)$$

for all gambles  $f$  on  $\Omega$ . A similar result for p-boxes on  $[0, 1]$  can be found in [41, Section 4.6.6] and [33, Theorem 2]. To see why this holds, note that  $\underline{P}_{\underline{F}, \overline{F}}$  is the lower envelope of the set  $\mathcal{M}(\underline{P}_{\underline{F}, \overline{F}})$  of linear previsions dominating  $\underline{P}_{\underline{F}, \overline{F}}$ , because  $\underline{P}_{\underline{F}, \overline{F}}$  is a coherent lower probability. Each of the linear previsions  $Q$  in  $\mathcal{M}(\underline{P}_{\underline{F}, \overline{F}})$  has a cumulative distribution function  $F_Q$ , and it is easy to see that  $F_Q \in \Phi(\underline{F}, \overline{F})$ . Conversely, any linear prevision whose cumulative distribution function  $F$  belongs to  $\Phi(\underline{F}, \overline{F})$  must belong to  $\mathcal{M}(\underline{P}_{\underline{F}, \overline{F}})$ . Therefore, the set  $\mathcal{M}(\underline{P}_{\underline{F}, \overline{F}})$  coincides with the set of all linear previsions whose cumulative distribution function belongs to  $\Phi(\underline{F}, \overline{F})$ , which establishes Eq. (1). We shall use this equation repeatedly in subsequent proofs. This

<sup>1</sup>The proof is virtually identical to the one given in [39, p. 93, Thm. 3.59].

<sup>2</sup>The natural extension of a cumulative distribution function  $F$  is simply understood to be the natural extension of the precise p-box  $(F, F)$ .

FIGURE 1. Examples of p-boxes on  $[0, 1]$ .

allows us moreover to give a sensitivity analysis interpretation to p-boxes: we can always regard them as a model for the imprecise knowledge of a cumulative distribution function.

We end this section with a trivial, yet very useful, approximation theorem:

**Theorem 2.** *Let  $\underline{P}$  be any coherent lower prevision defined on  $\mathcal{L}$ . The least conservative p-box  $(\underline{F}, \overline{F})$  on  $(\Omega, \leq)$  whose natural extension is dominated by  $\underline{P}$  is given by*

$$\underline{F}(x) = \underline{P}([0_\Omega, x]) \quad \overline{F}(x) = \overline{P}([0_\Omega, x])$$

for all  $x \in \Omega$ .

*Proof.* Obviously, the natural extension  $\underline{E}_{\underline{F}, \overline{F}}$  of the p-box, with  $\underline{F}$  and  $\overline{F}$  as above, is dominated by  $\underline{P}$ , because  $\underline{P}$  is an extension of  $\underline{P}_{\underline{F}, \overline{F}}$  (see for instance [39, p. 98, Prop. 4.7]).

Any other p-box  $(\underline{G}, \overline{G})$  whose natural extension is dominated by  $\underline{P}$  must satisfy:

$$\begin{aligned} \underline{G}(x) &= \underline{P}_{\underline{G}, \overline{G}}([0_\Omega, x]) = \underline{E}_{\underline{G}, \overline{G}}([0_\Omega, x]) \leq \underline{P}([0_\Omega, x]) = \underline{F}(x) \\ \overline{G}(x) &= 1 - \underline{P}_{\underline{G}, \overline{G}}((x, 1]) = \overline{E}_{\underline{G}, \overline{G}}([0_\Omega, x]) \geq \overline{P}([0_\Omega, x]) = \overline{F}(x) \end{aligned}$$

so  $(\underline{F}, \overline{F})$  is indeed the least conservative one.  $\square$

#### 4. NATURAL EXTENSION TO ALL EVENTS

The remainder of this paper is devoted to the natural extension  $\underline{E}_{\underline{F}, \overline{F}}$  of  $\underline{P}_{\underline{F}, \overline{F}}$ , and to various convenient expressions for it. We start by giving the form of the natural extension on the field of events generated by  $\mathcal{K}$ .

**4.1. Extension to the Field Generated by the Domain.** Let  $\mathcal{H}$  be the field of events generated by the domain  $\mathcal{K}$  of the p-box, i.e., events of the type

$$[0_\Omega, x_1] \cup (x_2, x_3] \cup \dots \cup (x_{2n}, x_{2n+1}]$$

for  $x_1 < x_2 < x_3 < \dots < x_{2n+1}$  in  $\Omega$  (if  $n$  is 0 we simply take this expression to be  $[0_\Omega, x_1]$ ) and

$$(x_2, x_3] \cup \dots \cup (x_{2n}, x_{2n+1}]$$

for  $x_2 < x_3 < \dots < x_{2n+1}$  in  $\Omega$ . Clearly, these events form a field: the union and intersection of any two events in  $\mathcal{H}$  is again in  $\mathcal{H}$ , and the complement of any event in  $\mathcal{H}$  also is again in  $\mathcal{H}$ .

To simplify the description of this field, and the expression of natural extension, we introduce an element  $0_\Omega-$  such that:<sup>3</sup>

$$0_\Omega- < x \text{ for all } x \in \Omega$$

$$F(0_\Omega-) = \underline{F}(0_\Omega-) = \overline{F}(0_\Omega-) = 0$$

In particular,  $(0_\Omega-, x] = [0_\Omega, x]$ . If we let  $\Omega^* = \Omega \cup \{0_\Omega-\}$ , then

$$\mathcal{H} = \{(x_0, x_1] \cup (x_2, x_3] \cup \dots \cup (x_{2n}, x_{2n+1}]: x_0 < x_1 < \dots < x_{2n+1} \text{ in } \Omega^*\}. \quad (2)$$

Since the procedure of natural extension is transitive, in order to calculate the natural extension of  $\underline{P}_{\underline{F}, \overline{F}}$  to all gambles, we shall first consider the extension from  $\mathcal{K}$  to  $\mathcal{H}$ , then the natural extension from  $\mathcal{H}$  to the set of all events, and finally the natural extension from the set of all events to the set of all gambles.

In the case of a precise p-box,  $\underline{P}_F$  has a unique extension to a finitely additive probability measure on  $\mathcal{H}$ .

**Proposition 3.**  *$\underline{E}_F$  restricted to  $\mathcal{H}$  is a finitely additive probability measure. Moreover, for any  $A \in \mathcal{H}$ , that is  $A = (x_0, x_1] \cup (x_2, x_3] \cup \dots \cup (x_{2n}, x_{2n+1}]$  with  $x_0 < x_1 < \dots < x_{2n+1}$  in  $\Omega^*$ , it holds that*

$$\underline{E}_F(A) = \sum_{k=0}^n (F(x_{2k+1}) - F(x_{2k})) \quad (3)$$

*Proof.* Because

$$\underline{E}_F([0_\Omega, x]) = \underline{P}_F([0_\Omega, x]) = F(x) \text{ and}$$

$$\overline{E}_F([0_\Omega, x]) = 1 - \underline{E}_F((x, 1_\Omega]) = 1 - \underline{P}_F((x, 1_\Omega]) = F(x)$$

for all  $x$ ,  $\underline{E}_F$  is linear on the linear space spanned by  $\{[0_\Omega, x]: x \in \Omega\}$  [39, p. 102, Prop. 4.18]. This linear space includes  $\mathcal{H}$ , which proves that  $\underline{E}_F$  is additive on  $\mathcal{H}$ , and consequently it is a finitely additive probability measure on  $\mathcal{H}$ . The expressions for  $\underline{E}_F(A)$  follow immediately.  $\square$

We now extend Proposition 3 to p-boxes.

**Proposition 4.** *For any  $A \in \mathcal{H}$ , that is  $A = (x_0, x_1] \cup (x_2, x_3] \cup \dots \cup (x_{2n}, x_{2n+1}]$  with  $x_0 < x_1 < \dots < x_{2n+1}$  in  $\Omega^*$ , it holds that  $\underline{E}_{\underline{F}, \overline{F}}(A) = \underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A)$ , where*

$$\underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A) = \sum_{k=0}^n \max\{0, \underline{F}(x_{2k+1}) - \overline{F}(x_{2k})\}. \quad (4)$$

*Proof.* We first show that  $\underline{E}_{\underline{F}, \overline{F}}(A) \leq \underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A)$ . Consider a cumulative distribution function  $F$  in  $\Phi(\underline{F}, \overline{F})$  which satisfies

$$F(x_{2k}) = \overline{F}(x_{2k})$$

$$F(x_{2k+1}) = \max\{\underline{F}(x_{2k+1}), \overline{F}(x_{2k})\}$$

<sup>3</sup>The cunning reader notes that one could avoid introducing  $0_\Omega-$  by imposing  $\underline{F}(0_\Omega) = \overline{F}(0_\Omega) = 0$ . However, this leads to an apparent loss of generality when linking p-boxes to other uncertainty models, and a slightly more complicated equation for the natural extension—therefore, for our purpose, it is simpler to stick to the most general formulation. ( $\underline{F}(1_\Omega) = \overline{F}(1_\Omega) = 1$  follows from coherence, so nothing is lost there.)

for all  $k = 0, \dots, n$ . Note that  $0 \leq F(x_0) \leq \dots \leq F(x_{2n+1}) \leq 1$ , so there is a cumulative distribution function satisfying the above equalities. Secondly, note that for each  $k = 0, \dots, 2n+1$ ,  $F$  satisfies  $\underline{F}(x_k) \leq F(x_k) \leq \overline{F}(x_k)$ . Hence, there is indeed a cumulative distribution function  $F$  in  $\Phi(\underline{F}, \overline{F})$  satisfying the above equalities. By Eq. (3),

$$\begin{aligned} \underline{E}_F(A) &= \sum_{k=0}^n \max\{\underline{F}(x_{2k+1}), \overline{F}(x_{2k})\} - \overline{F}(x_{2k}) \\ &= \sum_{k=0}^n \max\{0, \underline{F}(x_{2k+1}) - \overline{F}(x_{2k})\} = \underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A) \end{aligned}$$

with respect to  $F$ . Using Eq. (1), we deduce that  $\underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A) \geq \underline{E}_{\underline{F}, \overline{F}}(A)$ .

Next, we show that  $\underline{E}_{\underline{F}, \overline{F}}(A) \geq \underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A)$ . Let  $F$  be any cumulative distribution function in  $\Phi(\underline{F}, \overline{F})$ . Then,

$$\underline{E}_F((x_{2k}, x_{2k+1}]) = F(x_{2k+1}) - F(x_{2k}) \geq \max\{0, \underline{F}(x_{2k+1}) - \overline{F}(x_{2k})\}$$

since  $F(x_{2k+1}) \geq \underline{F}(x_{2k+1})$  and  $-F(x_{2k}) \geq -\overline{F}(x_{2k})$ . But, because  $\underline{E}_F$  is a finitely additive probability measure on  $\mathcal{H}$  (Proposition 3),  $\underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A) \leq \underline{E}_F(A)$ . This holds for any  $F$  in  $\Phi(\underline{F}, \overline{F})$ , and hence Eq. (1) implies that  $\underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A) \leq \underline{E}_{\underline{F}, \overline{F}}(A)$ .  $\square$

Note that it is possible to derive a closed expression for the upper prevision  $\overline{E}_{\underline{F}, \overline{F}}$  as well, similar to Eq. (4), however that expression is not as easy to work with. In practice, to calculate  $\overline{E}_{\underline{F}, \overline{F}}(A)$  for an event  $A \in \mathcal{H}$ , it is by far easiest first to calculate  $\underline{E}_{\underline{F}, \overline{F}}(A^c)$  using Eq. (4) (observe that also  $A^c \in \mathcal{H}$ ), and then to apply the conjugacy relation:

$$\overline{E}_{\underline{F}, \overline{F}}(A) = 1 - \underline{E}_{\underline{F}, \overline{F}}(A^c).$$

The lower probability  $\underline{P}_{\underline{F}, \overline{F}}$  on  $\mathcal{K}$  determined by the p-box usually does not have a *unique* coherent extension to the field  $\mathcal{H}$  (unless  $\underline{F} = \overline{F}$ ), as shown in the following example.

*Example 5.* Consider the distribution functions  $\underline{F}$  and  $\overline{F}$  given by  $\overline{F}(x) = x$  for  $x \in [0, 1]$  and  $\underline{F}(x) = 0$  if  $x \leq 0.5$ ,  $\underline{F}(x) = 2(x - 0.5)$  if  $x \geq 0.5$ .

From Proposition 3, both  $\underline{P}_{\underline{F}}$  and  $\underline{P}_{\overline{F}}$  have a unique extension to the field  $\mathcal{H}$ . Let us define  $\underline{P}_1$  on  $\mathcal{H}$  by  $\underline{P}_1(A) := \min\{\underline{P}_{\underline{F}}(A), \underline{P}_{\overline{F}}(A)\}$  for all  $A$ . Then  $\underline{P}_1$  is a coherent lower prevision on  $\mathcal{H}$ , and it is not difficult to show that  $\underline{P}_1 = \underline{P}_{\underline{F}, \overline{F}}(A)$  for any  $A \in \mathcal{K}$ .

Now, given the interval  $(0.5, 0.6]$ , we deduce from Proposition 3 that

$$\begin{aligned} \underline{P}_1((0.5, 0.6]) &= \min\{\underline{P}_{\underline{F}}((0.5, 0.6]), \underline{P}_{\overline{F}}((0.5, 0.6])\} \\ &= \min\{\underline{F}(0.6) - \underline{F}(0.5), \overline{F}(0.6) - \overline{F}(0.5)\} = \min\{0.2, 0.1\} = 0.1. \end{aligned}$$

However, it follows from Proposition 4 that  $\underline{E}_{\underline{F}, \overline{F}}((0.5, 0.6]) = \max\{0, \underline{F}(0.6) - \overline{F}(0.5)\} = \max\{0, 0.2 - 0.5\} = 0$ .

**4.2. Inner Measure.** The inner measure  $\underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}$  of the coherent lower probability  $\underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}$  defined in Eq. (4) coincides with  $\underline{E}_{\underline{F}, \overline{F}}$  on all events [41, Cor. 3.1.9, p. 127]:

$$\underline{E}_{\underline{F}, \overline{F}}(A) = \underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A) = \sup_{C \in \mathcal{H}, C \subseteq A} \underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(C), \quad (5)$$

The next example demonstrates that the choice of the preorder  $\leq$  can have a significant impact, even for the same  $\underline{F}$  and  $\overline{F}$ .



*Example 6.* Take  $\Omega = \{0, 1, 2, 3, 4\}$  and consider:

$$\underline{F}(x) = \overline{F}(x) = \begin{cases} 0 & \text{if } x < 2, \\ 1 & \text{otherwise.} \end{cases}$$

Consider the total preorder  $\leq_1$  defined by  $0 \simeq_1 1 <_1 2 \simeq_1 3 \simeq_1 4$  and the usual total ordering  $\leq_2$  defined by  $0 <_2 1 <_2 2 <_2 3 <_2 4$ . With  $\leq_1$ , we have for any event  $A \subseteq \Omega$  that

$$\underline{E}_{\underline{F}, \overline{F}}(A) = \begin{cases} 1 & \text{if } \{2, 3, 4\} \subseteq A, \\ 0 & \text{otherwise.} \end{cases}$$

using Eqs. (4) and (5). However, with  $\leq_2$ ,

$$\underline{E}_{\underline{F}, \overline{F}}(A) = \begin{cases} 1 & \text{if } 2 \in A, \\ 0 & \text{otherwise.} \end{cases}$$

For ease of notation, from now onwards, we denote  $\underline{E}_{\underline{F}, \overline{F}}$  by  $\underline{E}$  when no confusion about the cumulative distribution functions determining the p-box can arise.

In principle, the problem of natural extension to all events is solved: simply calculate the inner measure as in Eq. (5), using Eq. (4) to calculate  $\underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(C)$  for elements  $C$  in  $\mathcal{H}$ . However, the inner measure still involves calculating a supremum, which may be non-obvious. What we show next is that Eq. (4) can be extended to arbitrary events, by first taking the topological interior with respect to a very simple topology, followed by a (possibly infinite) sum over the so-called full components of this interior.

**4.3. The Partition Topology.** Consider the *partition topology* on  $\Omega$  generated by the equivalence classes  $[x]_{\simeq}$ , that is, the topology generated by

$$\tau := \{[x]_{\simeq} : x \in \Omega\}.$$

This topology is very similar to the discrete topology, except that it is not Hausdorff, unless  $\leq$  is anti-symmetric: if  $x \simeq y$  but  $x \neq y$  then  $x$  and  $y$  cannot be topologically separated, since every neighborhood of  $x$  is also a neighborhood of  $y$  and vice versa. If  $\leq$  is anti-symmetric (for example, the usual ordering on the reals is), then the partition topology reduces to the discrete topology, that is, every set is clopen (closed and open).

The open sets in this topology are all unions of equivalence classes (or, subsets of  $\Omega / \simeq$ , if you like). Hence, in this topology, every open set is also closed. In particular, every interval in  $(\Omega, \leq)$  is clopen.

The topological interior of a set  $A$  is given by the union of all equivalence classes contained in  $A$ :

$$\text{int}(A) = \bigcup \{[x]_{\simeq} : [x]_{\simeq} \subseteq A\} \quad (6)$$

and the topological closure is given by the union of all equivalence classes which intersect with  $A$ :

$$\text{cl}(A) = \bigcup \{[x]_{\simeq} : [x]_{\simeq} \cap A \neq \emptyset\}. \quad (7)$$

**Lemma 7.** For any subset  $A$  of  $\Omega$ ,  $\underline{E}(A) = \underline{E}(\text{int}(A))$  and  $\overline{E}(A) = \overline{E}(\text{cl}(A))$ .

*Proof.* Clearly  $\underline{E}(\text{int}(A)) \leq \underline{E}(A)$  because  $\text{int}(A) \subseteq A$ . If we can also show that  $\underline{E}(\text{int}(A)) \geq \underline{E}(A)$  the desired result is established.

Consider an element  $C$  of the field  $\mathcal{H}$  which is included in  $A$ . Since  $C$  is in particular an open set in the partition topology, it is a subset of  $\text{int}(A)$ . The monotonicity of  $\underline{E}$  implies that  $\underline{E}(C) \leq \underline{E}(\text{int}(A))$ . Consequently,

$$\underline{E}(A) = \underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(A) = \sup_{C \in \mathcal{H}, C \subseteq A} \underline{P}_{\underline{F}, \overline{F}}^{\mathcal{H}}(C) = \sup_{C \in \mathcal{H}, C \subseteq A} \underline{E}(C) \leq \underline{E}(\text{int}(A)).$$

The equality  $\overline{E}(A) = \overline{E}(\text{cl}(A))$  now follows from the fact that  $\text{cl}(A) = (\text{int}(A^c))^c$  and  $\overline{E}(A) = 1 - \underline{E}(A^c)$   $\square$

**4.4. Additivity on Full Components.** Next, we determine a constructive expression of the natural extension  $\underline{E}$  on the clopen subsets of  $\Omega$ .

**Definition 8.** [37, §4.4] A set  $S \subseteq \Omega$  is called *full* if  $[a, b] \subseteq S$  for any  $a \leq b$  in  $S$ .

What do these full sets look like? Obviously, any full set is clopen, as it must be a union of equivalence classes.

**Lemma 9.** *Every full set is clopen.*

*Proof.* Observe that  $[a, a] = [a]_{\simeq}$ , and apply Definition 8.  $\square$

Under an additional completeness assumption, the full sets are precisely the intervals.

**Lemma 10.** *If  $\Omega/\simeq$  is order complete, that is, if every subset of  $\Omega/\simeq$  has a supremum (minimal upper bound) and infimum (maximal lower bound), then every full set is an interval, that is, it can be written as  $[x, y]$ ,  $[x, y)$ ,  $(x, y]$ , or  $(x, y)$ , for some  $x, y$  in  $\Omega$ .*

*Proof.* Consider a full set  $S$  in  $\Omega$ . Since, by Lemma 9,  $S$  is clopen, we may consider  $S$  as a subset of  $\Omega/\simeq$ . So  $[x]_{\simeq} = \inf S$  and  $[y]_{\simeq} = \sup S$  exist, by the order completeness of  $\Omega/\simeq$ . Consider the case  $x \in S$  and  $y \notin S$ . Apply the definitions of  $\inf$  and  $\sup$  to establish that  $z \in S$  if and only if  $x \leq z < y$ . The other three cases are proven similarly.  $\square$

Note that  $\Omega/\simeq$  can be made order complete via the Dedekind completion [37, §4.34].

**Definition 11.** [37, §4.4] Given a clopen set  $A \subseteq \Omega$  and an element  $x$  of  $A$ , the *full component*  $C(x, A)$  of  $x$  in  $A$  is the largest full set  $S$  which satisfies  $x \in S \subseteq A$ .

**Lemma 12.** *The full components of any clopen set  $A$  form a partition of  $A$ .*

*Proof.* This is easily shown using a similar result for total orders given in [37, 4.4(a)].  $\square$

In the following theorem, we prove that the natural extension  $\underline{E}$  is additive on full components. Recall that the sum of a (possibly infinite) family  $(x_\lambda)_{\lambda \in \Lambda}$  of non-negative real numbers is defined as

$$\sum_{\lambda \in \Lambda} x_\lambda = \sup_{\substack{L \subseteq \Lambda \\ L \text{ finite}}} \sum_{\lambda \in L} x_\lambda$$

If the outcome of the above sum is a finite number, at most countably many of the  $x_\lambda$ 's are non-zero [37, 10.40].

**Theorem 13.** *Let  $B$  be a clopen subset of  $\Omega$ . Let  $(B_\lambda)_{\lambda \in \Lambda}$  be the full components of  $B$ , and let  $(C_\lambda)_{\lambda \in \Lambda'}$  be the full components of  $B^c$ . Then*

$$\begin{aligned} \underline{E}(B) &= \sum_{\lambda \in \Lambda} \underline{E}(B_\lambda) \\ \overline{E}(B) &= 1 - \sum_{\lambda \in \Lambda'} \underline{E}(C_\lambda) \end{aligned}$$

*Proof.* Since  $(B_\lambda)_{\lambda \in \Lambda}$  is a partition of  $B$ , and because the functional  $\underline{E}$  is monotone and super-additive,

$$\underline{E}(B) = \underline{E}(\cup_{\lambda \in \Lambda} B_\lambda) \geq \underline{E}(\cup_{\lambda \in L} B_\lambda) \geq \sum_{\lambda \in L} \underline{E}(B_\lambda)$$

for every finite subset  $L$  of  $\Lambda$ , and consequently  $\underline{E}(B) \geq \sum_{\lambda \in \Lambda} \underline{E}(B_\lambda)$ . We are left to show that also  $\underline{E}(B) \leq \sum_{\lambda \in \Lambda} \underline{E}(B_\lambda)$ .

Let  $\epsilon > 0$ . By Eq. (5), the inner measure coincides with the natural extension, so there is a  $C$  in  $\mathcal{H}$  such that  $C \subseteq B$  and  $\underline{E}(B) \leq \underline{E}(C) + \epsilon$ . From Eq. (2), we can write

$$C = (x_0, x_1] \cup (x_2, x_3] \cup \cdots \cup (x_{2n}, x_{2n+1}]$$

for some  $x_0 < x_1 < \cdots < x_{2n+1}$  in  $\Omega^*$ . Let us denote  $C_k := (x_{2k}, x_{2k+1}]$  for  $k = 0, \dots, n$ . It is easily established that  $C_0, \dots, C_n$  are the full components of  $C$ .

Applying Proposition 4 twice, and using the fact that each full component  $C_k$  of  $C$  must be a subset of exactly one full component  $B_\lambda$  of  $B$  (this is a consequence of  $C \subseteq B$ ), we find that

$$\underline{E}(C) = \sum_{k=0}^n \underline{E}(C_k) = \sum_{\lambda \in \Lambda} \underline{E}(\cup_{k: C_k \subseteq B_\lambda} C_k)$$

(note that the union is  $\emptyset$  for those  $\lambda$  where no  $C_k \subseteq B_\lambda$ —so only a finite number of terms can be non-zero in the latter sum) and consequently

$$\underline{E}(B) \leq \underline{E}(C) + \epsilon = \sum_{\lambda \in \Lambda} \underline{E}(\cup_{k: C_k \subseteq B_\lambda} C_k) + \epsilon \leq \sum_{\lambda \in \Lambda} \underline{E}(B_\lambda) + \epsilon,$$

taking into account the monotonicity of  $\underline{E}$ . As this holds for any  $\epsilon > 0$ , we arrive at the desired inequality.

The expression for  $\overline{E}$  simply follows from the conjugacy relation  $\overline{E}(B) = 1 - \underline{E}(B^c)$ , once noted that  $B^c$  is clopen as well.  $\square$

In other words, the natural extension  $\underline{E}$  of a p-box is *arbitrarily additive on full components*. In particular, interestingly, it is  $\sigma$ -additive on full components (but obviously not additive, let alone  $\sigma$ -additive, on arbitrary events).

*Example 14.* Additivity on full components is not sufficient for a lower probability to be equivalent to a p-box, even in the finite case. For example, consider  $\Omega = \{1, 2, 3\}$  with the usual ordering, so  $\mathcal{H} = \wp(\Omega)$ . Let  $\underline{P}$  be the lower probability defined by

$$\underline{P}(\{1\}) = \underline{P}(\{2\}) = \underline{P}(\{3\}) = 0.1$$

It can be checked that  $\underline{P}$  is coherent, and that the natural extension  $\underline{E}$  of  $\underline{P}$  is the lower envelope of the probability mass functions  $(0.8, 0.1, 0.1)$ ,  $(0.1, 0.8, 0.1)$ , and  $(0.1, 0.1, 0.8)$ . Moreover,  $\underline{E}$  is additive on full components, because

$$\underline{E}(\{1, 3\}) = \underline{E}(\{1\}) + \underline{E}(\{3\})$$

(every other subset of  $\Omega$  is already full). However,

$$\underline{E}(\{2\}) \neq \max\{0, \underline{E}(\{1, 2\}) - \overline{E}(\{1\})\}$$

because  $\underline{E}(\{1, 2\}) = 0.2$  and  $\overline{E}(\{1\}) = 0.8$ . This shows that  $\underline{E}$  violates Proposition 4 and as a consequence it is not the natural extension to events of a p-box.

**4.5. Summary.** Let us summarize all results so far, and explain how, in practice,  $\underline{E}(A)$  and  $\overline{E}(A)$  of an arbitrary event  $A$  can be calculated.

Proposition 4 gave the natural extension to the field  $\mathcal{H}$ ; we are now in a position to generalize it to all events, at least when  $\Omega/\simeq$  is order complete.

Indeed, consider an arbitrary event  $A$ . By Lemma 7, it suffices to calculate the natural extension of  $\text{int}(A)$  or  $\text{cl}(A)$ . Calculating the interior or closure with respect to the partition topology will usually be trivial (see examples further on), and moreover, the topological interior or closure of a set is always clopen, so now we only need to know the natural extension of clopen sets.

Now, by Theorem 13, it follows that we only need to calculate the natural extension of the full components  $(B_\lambda)_{\lambda \in \Lambda}$  of  $\text{int}(A)$  or the full components  $(C_\lambda)_{\lambda \in \Lambda}$  of  $\text{cl}(A)^c = \text{int}(A^c)$ —note that

each of these full components is also clopen. Also, finding the full components will often be a trivial operation—commonly, there will only be a few.

But, by Lemma 10, if, in addition,  $\Omega/\simeq$  is order complete, then each full component is an interval. And for intervals, we immediately infer from Proposition 4 and Eq. (5) that:

$$\underline{E}((x, y]) = \max\{0, \underline{F}(y) - \overline{F}(x)\} \quad (8a)$$

$$\underline{E}((x, y)) = \max\{0, \underline{F}(y-) - \overline{F}(x)\} \quad (8b)$$

$$\underline{E}([x, y]) = \begin{cases} \max\{0, \underline{F}(y) - \overline{F}(x)\} & \text{if } x \text{ has no immediate predecessor} \\ \max\{0, \underline{F}(y) - \overline{F}(x-)\} & \text{if } x \text{ has an immediate predecessor} \end{cases} \quad (8c)$$

$$\underline{E}([x, y)) = \begin{cases} \max\{0, \underline{F}(y-) - \overline{F}(x)\} & \text{if } x \text{ has no immediate predecessor} \\ \max\{0, \underline{F}(y-) - \overline{F}(x-)\} & \text{if } x \text{ has an immediate predecessor} \end{cases} \quad (8d)$$

for any  $x < y$  in  $\Omega$ ,<sup>4</sup> where  $\underline{F}(y-)$  denotes  $\sup_{z < y} \underline{F}(z)$  and similarly for  $\overline{F}(x-)$ . The equalities hold because, if  $x < y$  in  $\Omega$ , and  $x-$  is an immediate predecessor of  $x$ , then  $[x, y] = (x-, y]$  and  $[x, y) = (x-, y)$ . Recall also that  $\underline{F}(0_\Omega-) = \overline{F}(0_\Omega-) = 0$  by convention. If  $\Omega/\simeq$  is finite, then one can think of  $z-$  as the immediate predecessor of  $z$  in the quotient space  $\Omega/\simeq$  for any  $z \in \Omega$ ;

In other words, we have a simple constructive means of calculating the natural extension of any event.

**4.6. Special Cases.** The above equations hold for any  $(\Omega, \leq)$  with order complete quotient space. In most cases in practice, either

- $\Omega/\simeq$  is finite, or
- $\Omega/\simeq$  is connected meaning that for any two elements  $x < y$  in  $\Omega$  there is a  $z$  in  $\Omega$  such that  $x < z < y$ ,<sup>5</sup> (this is the case for instance when  $\Omega$  is a closed interval in  $\mathbb{R}$  and  $\leq$  is the usual ordering of reals)

Moreover, if  $\Omega/\simeq$  is connected, then  $\underline{F}$  will commonly satisfy  $\underline{F}(y-) = \underline{F}(y)$  for all  $y$  in  $\Omega$ . For example, in case  $\Omega$  is a closed interval in  $\mathbb{R}$ , this happens precisely when  $\underline{F}(0) = 0$  and  $\underline{F}$  is left-continuous in the usual sense.

Obviously, if  $\Omega/\simeq$  is finite, then every element of  $\Omega$  has an immediate predecessor (remember, we take the immediate predecessor of  $0_\Omega$  to be  $0_\Omega-$ ), and if  $\Omega/\simeq$  is connected, then no element except  $0_\Omega$  has an immediate predecessor.

By Lemma 7, Theorem 13, and Eqs. (8), we conclude:

**Corollary 15.** *If  $\Omega/\simeq$  is finite, then every full set  $B$  is of the form  $[a, b]$ , and*

$$\begin{aligned} \underline{E}(A) &= \sum_{\lambda \in \Lambda} \max\{0, \underline{F}(b_\lambda) - \overline{F}(a_\lambda-)\} \\ \overline{E}(A) &= 1 - \sum_{\lambda \in \Lambda'} \max\{0, \underline{F}(b'_\lambda) - \overline{F}(a'_\lambda-)\} \end{aligned}$$

where  $([a_\lambda, b_\lambda])_{\lambda \in \Lambda}$  are the full components of  $\text{int}(A)$ , and  $([a'_\lambda, b'_\lambda])_{\lambda \in \Lambda'}$  are the full components of  $\text{int}(A^c) = \text{cl}(A)^c$ .

<sup>4</sup>In case  $x = 0_\Omega$ , evidently,  $0_\Omega-$  is the immediate predecessor.

<sup>5</sup>This terminology stems from the fact that, in this case,  $\Omega/\simeq$  is connected with respect to the order topology [37, §15.46(6)].

**Corollary 16.** *If  $\Omega/\simeq$  is order complete and connected, and  $\underline{F}(y-) = \underline{F}(y)$  for all  $y$  in  $\Omega$ , then*

$$\begin{aligned}\underline{E}(A) &= \sum_{\lambda \in \Lambda} \max\{0, \underline{F}(\sup B_\lambda) - \overline{F}(\inf B_\lambda)\} \\ \overline{E}(A) &= 1 - \sum_{\lambda \in \Lambda'} \max\{0, \underline{F}(\sup C_\lambda) - \overline{F}(\inf C_\lambda)\}\end{aligned}$$

where  $(B_\lambda)_{\lambda \in \Lambda}$  are the full components of  $\text{int}(A)$  and  $(C_\lambda)_{\lambda \in \Lambda'}$  are the full components of  $\text{int}(A^c) = \text{cl}(A)^c$ .

Beware of  $\underline{F}(0_\Omega) = \underline{F}(0_\Omega-) = 0$  in the last corollary.

**4.7. Example.** Let's investigate a particular type of p-boxes on the unit square  $[0, 1]^2$ . First, we must specify a preorder on  $\Omega$ . A natural yet naive way of doing so is, for instance, saying that  $(x_1, y_1) \leq (x_2, y_2)$  whenever

$$x_1 + y_1 \leq x_2 + y_2$$

Consider a p-box  $(\underline{F}, \overline{F})$  on  $([0, 1]^2, \leq)$ . Since  $\underline{F}$  is required to be non-decreasing with respect to  $\leq$ , it follows that  $\underline{F}(x, y)$  is constant on elements of  $[0, 1]^2/\simeq$ , which means that  $\underline{F}(x_1, y_1) = \underline{F}(x_2, y_2)$  whenever  $x_1 + y_1 = x_2 + y_2$ . Thus, we may think of  $\underline{F}(x, y)$  as a function of a single variable  $z = x + y$ , and we write  $\underline{F}(z)$ . Similarly, we write  $\overline{F}(z)$ .

Our definition of  $\leq$  means that our p-box specifies bounds on the probability of right-angled triangles (restricted to  $[0, 1]^2$ ) whose hypotenuses are orthogonal to the diagonal:

$$\underline{E}(z) \leq p(\{(x, y) \in [0, 1]^2 : x + y \leq z\}) \leq \overline{F}(z) \quad (9)$$

Observe that the p-box is given directly on the two-dimensional product space, without the need to define marginal p-boxes for each dimension. The base  $\tau$  for our partition topology is given by

$$\tau = \{\{(x, y) \in [0, 1]^2 : x + y = z\} : z \in [0, 2]\}$$

For example, the topological interior of a rectangle  $A = [a, b] \times [c, d]$  is empty, unless  $a = c = 0$  or  $b = d = 1$ , because in all other cases, no element of  $\tau$  is a subset of  $A$ . In the cases where  $a = c = 0$  and  $\min\{b, d\} < 1$ , or  $\max\{a, c\} > 0$  and  $b = d = 1$  (if  $a = c = 0$  and  $b = d = 1$  then the interior is  $\Omega$ ), respectively, we have:

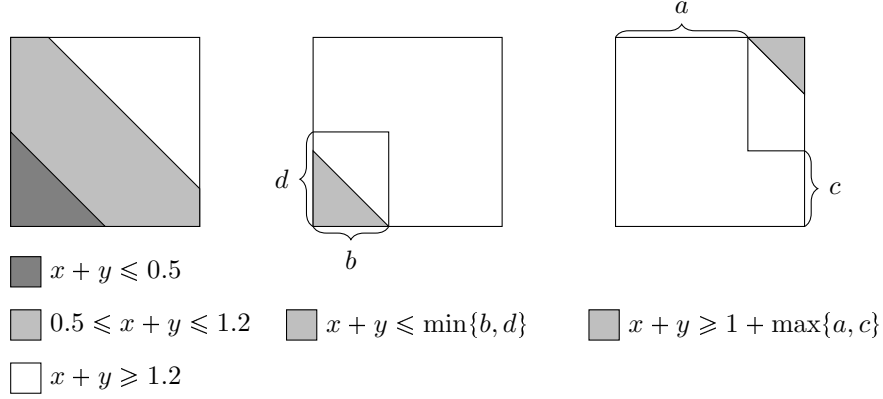
$$\begin{aligned}\text{int}([0, b] \times [0, d]) &= \{(x, y) \in [0, 1]^2 : x + y \leq \min\{b, d\}\} \\ \text{int}([a, 1] \times [c, 1]) &= \{(x, y) \in [0, 1]^2 : x + y \geq 1 + \max\{a, c\}\}\end{aligned}$$

Consequently,  $\underline{E}(A) = 0$  for all rectangles  $A$ , except for

$$\begin{aligned}\underline{E}([0, b] \times [0, d]) &= \underline{F}(\min\{b, d\}) \\ \underline{E}([a, 1] \times [c, 1]) &= 1 - \overline{F}(1 + \max\{a, c\})\end{aligned}$$

Fig. 2 illustrates the situation. So, for the purpose of making inferences about the lower probability of events that are rectangles, the ordering  $\leq$  was obviously poorly chosen. In general, *one should choose  $\leq$  in a way that  $\Omega/\simeq$  contains good approximations for all events of interest.*

For example, in the case of rectangles, we could for instance discretize  $[0, 1]^2$  into smaller squares, and impose some ordering on these squares. Of course, it may not be entirely obvious how to interpret the lower and upper cumulative distribution functions on such discretized space, since there is no natural ordering on such discretization. Another strategy would be to start from a reference point (e.g., an elicited modal value) and then to choose the ordering  $\leq$  such that intervals correspond to concentric regions of interests around the reference point. Again, all of this is possible because our theory concerns p-boxes on arbitrary totally preordered spaces, and is not limited to the real line with its natural ordering. More realistic examples in which such concentric regions are used are given in Section 8.

FIGURE 2. Shape of intervals induced by  $\leq$ , and calculation of the topological interior.

## 5. NATURAL EXTENSION TO ALL GAMBLES

Next, we establish that p-boxes are completely monotone, and that therefore their natural extension to all gambles can be expressed as a Choquet integral. We further simplify the calculation of this Choquet integral via the lower and upper oscillation of gambles with respect to the partition topology introduced earlier.

**5.1. Complete Monotonicity.** As shown in [33, Section 3.1], the natural extension  $\underline{E}_F$  of a distribution function  $F$  on  $[0, 1]$  is completely monotone. It is fairly easy to generalise this result to distribution functions on an arbitrary totally preordered space  $\Omega$ .<sup>6</sup> However, given this, and Eq. (1), we cannot immediately deduce the complete monotonicity of  $\underline{E}_{F, \overline{F}}$ , because the lower envelope of a set of completely monotone lower previsions is not necessarily completely monotone. We prove next that such an envelope is indeed completely monotone in the case of p-boxes. This is an improvement with respect to previous results [19], where the relation between p-boxes and complete monotonicity was established for finite spaces.

Let  $\underline{P}_{\underline{E}, \overline{F}}^{\mathcal{H}}$  denote the restriction of  $\underline{E}_{F, \overline{F}}$  to  $\mathcal{H}$ , given by Proposition 4.

**Theorem 17.**  $\underline{P}_{\underline{E}, \overline{F}}^{\mathcal{H}}$  is a completely monotone coherent lower probability.

*Proof.* Clearly,  $\underline{P}_{\underline{E}, \overline{F}}^{\mathcal{H}}$  is coherent as it is the natural extension to  $\mathcal{H}$  of a coherent lower probability  $\underline{P}_{F, \overline{F}}$  [41, p. 123, 3.1.2(a)]. To prove that it is completely monotone, we must establish that for all  $p \in \mathbb{N}$ ,  $2 \leq p \leq n$ , and all  $A_1, \dots, A_p$  in  $\mathcal{H}$ :

$$\underline{P}_{\underline{E}, \overline{F}}^{\mathcal{H}}\left(\bigcup_{i=1}^p A_i\right) \geq \sum_{\emptyset \neq I \subseteq \{1, \dots, p\}} (-1)^{|I|+1} \underline{P}_{\underline{E}, \overline{F}}^{\mathcal{H}}\left(\bigcap_{i \in I} A_i\right). \quad (10)$$

For any  $p \in \mathbb{N}$ ,  $2 \leq p \leq n$ , and any  $A_1, \dots, A_p$  in  $\mathcal{H}$ , consider the finite field generated by  $A_1, \dots, A_p$ . Let  $\underline{Q}$  denote the restriction of  $\underline{P}_{\underline{E}, \overline{F}}^{\mathcal{H}}$  to this finite field. By [19, Sec. 3],  $\underline{Q}$  is completely monotone on this finite field. In particular, Eq. (10) is satisfied. But this means that Eq. (10) is satisfied for all  $p \in \mathbb{N}$ ,  $2 \leq p \leq n$ , and all  $A_1, \dots, A_p$  in  $\mathcal{H}$ , which establishes the theorem.  $\square$

<sup>6</sup>Indeed, by [9, Thm. 5 & Thm. 9], the natural extension of any finitely additive probability on a field is completely monotone.

**5.2. Choquet Integral Representation.** Complete monotonicity allows us to characterise the natural extension on all gambles, as we show in the following theorem:

**Theorem 18.** *The natural extension  $\underline{E}$  of  $\underline{P}_{\underline{E}, \overline{F}}$  is given by the Choquet integral*

$$\underline{E}(f) = \inf f + \int_{\inf f}^{\sup f} \underline{E}(\{f \geq t\}) dt$$

for every gamble  $f$ . Moreover,  $\underline{E}$  is completely monotone on all gambles. Similarly,

$$\overline{E}(f) = \inf f + \int_{\inf f}^{\sup f} \overline{E}(\{f \geq t\}) dt.$$

*Proof.* Immediate from Theorem 17 once observed that  $\underline{E} = \underline{P}_{\underline{E}, \overline{F}}^{\mathcal{H}}$  on all events, and [9, Theorems 8 and 9]. The latter two theorems state that:

- Given a coherent  $n$ -monotone ( $n \geq 2$ ) lower probability  $\underline{P}$  defined on a field (here,  $\mathcal{H}$ ), its natural extension to all gambles is given by its Choquet integral  $(C) \int \cdot d\underline{P}_*$
- If a coherent lower probability  $\underline{P}$  defined on a field is  $n$ -monotone ( $n \geq 2$ ), then its natural extension to all gambles is  $n$ -monotone.

□

**5.3. Lower and Upper Oscillation.** By Lemma 7, to turn Theorem 18 in an effective algorithm, we must calculate  $\text{int}(\{f \geq t\})$  for every  $t$ . Fortunately, there is a very simple way to do this.

For any gamble  $f$  on  $\Omega$  and any topological base  $\tau$ , define its *lower oscillation* as the gamble

$$\underline{\text{osc}}(f)(x) = \sup_{C \in \tau: x \in C} \inf_{y \in C} f(y)$$

For the partition topology which we introduced earlier, this simplifies to

$$\underline{\text{osc}}(f)(x) = \inf_{y \in [x]_{\simeq}} f(y) \quad (11)$$

The upper oscillation is:

$$\overline{\text{osc}}(f)(x) = -\underline{\text{osc}}(-f)(x) = \sup_{y \in [x]_{\simeq}} f(y) \quad (12)$$

For a subset  $A$  of  $\Omega$ , we deduce from the above definition and from Eq. (6) that the lower oscillation of  $I_A$  is  $I_{\text{int}(A)}$ , so the lower oscillation is the natural generalisation of the topological interior to gambles. Similarly, we see from Eq. (7) that the upper oscillation of  $I_A$  is  $I_{\text{cl}(A)}$ .

**Proposition 19.** *For any gamble  $f$  on  $\Omega$ ,*

$$\text{int}(\{f \geq t\}) = \{\underline{\text{osc}}(f) \geq t\} \quad (13a)$$

$$\text{cl}(\{f \geq t\}) = \{\overline{\text{osc}}(f) \geq t\} \quad (13b)$$

so, in particular,

$$\underline{E}(f) = \inf \underline{\text{osc}}(f) + \int_{\inf \underline{\text{osc}}(f)}^{\sup \underline{\text{osc}}(f)} \underline{E}(\{\underline{\text{osc}}(f) \geq t\}) dt = \underline{E}(\underline{\text{osc}}(f)) \quad (14a)$$

$$\overline{E}(f) = \inf \overline{\text{osc}}(f) + \int_{\inf \overline{\text{osc}}(f)}^{\sup \overline{\text{osc}}(f)} \overline{E}(\{\overline{\text{osc}}(f) \geq t\}) dt = \overline{E}(\overline{\text{osc}}(f)) \quad (14b)$$

*Proof.* Eqs. (13) are easily established using the definitions of interior and closure, and lower and upper oscillation. For example,

$$x \in \text{int}(\{f \geq t\})$$

if and only if there is a  $C$  in  $\tau$  such that  $x \in C$  and

$$C \subseteq \{f \geq t\}$$

But, for our choice of  $\tau$ , necessarily  $C = [x]_{\simeq}$  if  $x \in C \in \tau$ , so the above holds if and only if

$$\forall y \in [x]_{\simeq} : f(y) \geq t$$

And this holds if and only if

$$\underline{\text{osc}}(f)(x) \geq t$$

where we used the definition of  $\underline{\text{osc}}(f)$ .

Let us prove Eqs. (14). It follows from Eq. (11) that  $f \geq \underline{\text{osc}}(f)$ , and as a consequence  $\underline{E}(f) \geq \underline{E}(\underline{\text{osc}}(f))$ . We are left to prove that  $\underline{E}(f) \leq \underline{E}(\underline{\text{osc}}(f))$ .

Indeed, using Lemma 7, and Eqs. (13),

$$\underline{E}(f) = \inf f + \int_{\inf f}^{\sup f} \underline{E}(\{f \geq t\}) dt = \inf f + \int_{\inf f}^{\sup f} \underline{E}(\{\underline{\text{osc}}(f) \geq t\}) dt$$

and since obviously, by Eq. (11),  $\inf f = \inf \underline{\text{osc}}(f)$ ,

$$= \inf \underline{\text{osc}}(f) + \int_{\inf \underline{\text{osc}}(f)}^{\sup f} \underline{E}(\{\underline{\text{osc}}(f) \geq t\}) dt$$

and since  $\sup f \geq \sup \underline{\text{osc}}(f)$ , using the usual properties of the Riemann integral,

$$\begin{aligned} &= \inf \underline{\text{osc}}(f) + \int_{\inf \underline{\text{osc}}(f)}^{\sup \underline{\text{osc}}(f)} \underline{E}(\{\underline{\text{osc}}(f) \geq t\}) dt \\ &\quad + \int_{\sup \underline{\text{osc}}(f)}^{\sup f} \underline{E}(\{\underline{\text{osc}}(f) \geq t\}) dt \end{aligned}$$

Now use the fact that  $\{\underline{\text{osc}}(f) \geq t\} = \emptyset$  for  $t \in (\sup \underline{\text{osc}}(f), \sup f]$ , so the last term is zero.  $\square$

Concluding, to calculate the natural extension of any gamble, in practice, we must simply determine the full components of the cut sets of its lower or upper oscillation, and calculate a simple Riemann integral of a monotonic function.

Examples will be given in Section 8.

## 6. P-BOXES WHOSE PREORDERS ARE INDUCED BY A REAL-VALUED FUNCTION

In practice, the most convenient way to specify a preorder  $\leq$  on  $\Omega$  such that  $\Omega/\simeq$  is order complete and connected is by means of a bounded real-valued function  $Z: \Omega \rightarrow \mathbb{R}$ . For instance, in the example in Section 4.7, we used  $Z(x, y) = x + y$ . Also see [2] and [28].

Let us assume from now onwards that  $Z$  is a surjective mapping from  $\Omega$  to  $[0, 1]$ .

For any  $x$  and  $y$  in  $\Omega$ , define  $x \leq y$  whenever  $Z(x) \leq Z(y)$ . Because  $Z$  is surjective,  $\Omega/\simeq$  is order complete and connected. In particular,  $\Omega$  has a smallest and largest element, for which  $Z(0_\Omega) = 0$  and  $Z(1_\Omega) = 1$ . Moreover, we can think of any cumulative distribution function on  $(\Omega, \leq)$  as a function over a single variable  $z \in [0, 1]$ . Consequently, we can think of any p-box on  $(\Omega, \leq)$  as a p-box on  $([0, 1], \leq)$ . In particular, for any subset  $I$  of  $[0, 1]$  we write  $\underline{E}(I)$  for  $\underline{E}(Z^{-1}(I))$ . For example, for  $a, b$  in  $[0, 1]$ , and  $A = Z^{-1}((a, b]) \subseteq \Omega$ , we have that

$$\underline{E}(A) = \underline{E}((a, b]) = \max\{0, \underline{E}(a) - \overline{F}(b)\}$$



by Proposition 4. Similar expressions for other types of intervals follow from Eqs. (8).

The topological interior and closure can be related to the so-called *lower and upper inverse* of  $Z^{-1}$ . Indeed, consider the multi-valued mapping  $\Gamma := Z^{-1}: [0, 1] \rightarrow \wp(\Omega)$ . Because for every  $x$  in  $\Omega$ , it holds that  $[x]_{\simeq} = \Gamma(Z(x))$ , it follows that, for any subset  $A$  of  $\Omega$ ,  $\text{int}(A) = \Gamma(\Gamma_*(A))$ , and  $\text{cl}(A) = \Gamma(\Gamma^*(A))$ , where  $\Gamma_*$  and  $\Gamma^*$  denote the lower and upper inverse of  $\Gamma$  respectively, that is

$$\begin{aligned}\Gamma_*(A) &= \{z \in [0, 1] : \Gamma(z) \subseteq A\}, \text{ and} \\ \Gamma^*(A) &= \{z \in [0, 1] : \Gamma(z) \cap A \neq \emptyset\}\end{aligned}$$

(see for instance [16]).<sup>7</sup>

**Theorem 20.** *Let  $A$  be an arbitrary subset of  $\Omega$ . Then*

$$\begin{aligned}\underline{E}(A) &= \sum_{\lambda \in \Lambda} \underline{E}(I_\lambda) \\ \overline{E}(A) &= 1 - \sum_{\lambda \in \Lambda'} \underline{E}(J_\lambda)\end{aligned}$$

where  $(I_\lambda)_{\lambda \in \Lambda}$  are the full components of  $Z(\text{int}(A)) = \Gamma_*(A)$  and  $(J_\lambda)_{\lambda \in \Lambda'}$  are the full components of  $Z(\text{int}(A^c)) = Z(\text{cl}(A)^c) = \Gamma_*(A^c)$ .

If, in addition,  $\underline{E}$  is left-continuous as a function of  $z \in [0, 1]$  and  $\underline{E}(0) = 0$ , then

$$\begin{aligned}\underline{E}(A) &= \sum_{\lambda \in \Lambda} \max\{0, \underline{E}(\sup I_\lambda) - \overline{F}(\inf I_\lambda)\} \\ \overline{E}(A) &= 1 - \sum_{\lambda \in \Lambda'} \max\{0, \underline{E}(\sup J_\lambda) - \overline{F}(\inf J_\lambda)\}\end{aligned}$$

*Proof.* Indeed, by Corollary 16,

$$\underline{E}(A) = \underline{E}(\text{int}(A)) = \sum_{\lambda \in \Lambda} \underline{E}(B_\lambda)$$

where  $(B_\lambda)_{\lambda \in \Lambda}$  are the full components of  $\text{int}(A)$ . So, the result is established if we can show that  $(Z^{-1}(I_\lambda))_{\lambda \in \Lambda}$  are the full components of  $\text{int}(A)$ .

Obviously, since  $(I_\lambda)_{\lambda \in \Lambda}$  partitions  $Z(\text{int}(A))$ , it follows that  $(Z^{-1}(I_\lambda))_{\lambda \in \Lambda}$  partitions

$$\bigcup_{\lambda \in \Lambda} Z^{-1}(I_\lambda) = Z^{-1}(Z(\text{int}(A))) = \text{int}(A)$$

where the latter equality follows from the fact that  $\text{int}(A)$  is clopen, i.e., is a union of equivalence classes.

We are left to prove that each set  $Z^{-1}(I_\lambda)$  is a full component. Clearly,  $Z^{-1}(I_\lambda)$  is full: for any two  $x$  and  $y$  in  $Z^{-1}(I_\lambda)$ , it holds that

$$[x, y] = \{v \in \Omega : Z(x) \leq Z(v) \leq Z(y)\} = Z^{-1}([Z(x), Z(y)]) \subseteq Z^{-1}(I_\lambda)$$

where we used the fact that  $[Z(x), Z(y)] \subseteq I_\lambda$  in the last step.

Consider any  $x \in Z^{-1}(I_\lambda)$ . The desired result is established if we can show that  $Z^{-1}(I_\lambda)$  is the largest full set  $S$  which satisfies  $x \in S \subseteq \text{int}(A)$ .

Suppose  $S$  is larger, that is,  $S$  is full,  $x \in S \subseteq \text{int}(A)$ , and  $Z^{-1}(I_\lambda) \subsetneq S$ . Since both sets are clopen, it must be that there is some  $y \in S$  such that  $[y]_{\simeq} \cap Z^{-1}(I_\lambda) = \emptyset$ . But this implies that

$$I_\lambda = Z(Z^{-1}(I_\lambda)) \subsetneq Z(S)$$

<sup>7</sup>We follow the terminology in [34] and [22], among others. Beware that  $\Gamma_*$  and  $\Gamma^*$  are sometimes called *upper* and *lower* inverse instead [40, 3], or *strong* and *weak* inverse [14].

because  $Z(y)$  belongs to  $Z(S)$  but not to  $Z(Z^{-1}(I_\lambda))$ . But, this would mean that  $I_\lambda$  is not a full component of  $Z(\text{int}(A))$ —a contradiction. So,  $Z^{-1}(I_\lambda)$  must be the largest full set  $S$  which satisfies  $x \in S \subseteq \text{int}(A)$ .  $\square$

Regarding gambles, note that the lower oscillation is constant on equivalence classes (this follows immediately from its definition). Hence, we may consider  $\underline{osc}(f)$  for a gamble  $f$  on  $\Omega$  as a function of  $z \in [0, 1]$ , and we can use Proposition 19 to write:

**Proposition 21.** *For any gamble  $f$  on  $\Omega$ ,*

$$\begin{aligned} \underline{E}(f) &= \inf \underline{osc}(f) + \int_{\inf \underline{osc}(f)}^{\sup \underline{osc}(f)} \underline{E}(\{z : \underline{osc}(f)(z) \geq t\}) dt \\ \overline{E}(f) &= \inf \overline{osc}(f) + \int_{\inf \overline{osc}(f)}^{\sup \overline{osc}(f)} \overline{E}(\{z : \overline{osc}(f)(z) \geq t\}) dt \end{aligned}$$

## 7. CONSTRUCTING MULTIVARIATE P-BOXES FROM MARGINALS

In this section, we construct a multivariate p-box from marginal coherent lower previsions under arbitrary rules of combination. As special cases, we derive expressions for the joint,

- (i) either without any assumptions about dependence or independence between variables, that is, using the Fréchet-Hoeffding bounds [29],
- (ii) or assuming epistemic independence between all variables, that is, using the factorization property [7].

We also derive Williamson and Downs's [45] probabilistic arithmetic as a special case of our framework.

Specifically, consider  $n$  variables  $X_1, \dots, X_n$  assuming values in  $\mathcal{X}_1, \dots, \mathcal{X}_n$ , and assume that marginal lower previsions  $\underline{P}_1, \dots, \underline{P}_n$ , are given for each variable—each of these could be the natural extension of a p-box, although we do not require this. So, each  $\underline{P}_i$  is a coherent lower prevision on  $\mathcal{L}(\mathcal{X}_i)$ .

**7.1. Multivariate P-Boxes.** The first step in constructing our multivariate p-box is to define a mapping  $Z$  to induce a preorder  $\leq$  on  $\Omega = \mathcal{X}_1 \times \dots \times \mathcal{X}_n$ . The following choice works perfectly for our purpose:

$$Z(x_1, \dots, x_n) = \max_{i=1}^n Z_i(x_i)$$

where each  $Z_i$  is a surjective mapping from  $\mathcal{X}_i$  to  $[0, 1]$  and hence, also induces an marginal preorder  $\leq_i$  on  $\mathcal{X}_i$ . Each  $\underline{P}_i$  can be approximated by a p-box  $(\underline{F}_i, \overline{F}_i)$  on  $(\mathcal{X}_i, \leq_i)$ , defined by

$$\underline{F}_i(z) = \underline{P}_i(Z_i^{-1}([0, z])) \quad \overline{F}_i(z) = \overline{P}_i(Z_i^{-1}([0, z]))$$

This approximation is the best possible one, by Theorem 2.

Beware that even though different choices of  $Z_i$  may induce the same total preorder  $\leq_i$ , they might lead to a different total preorder  $\leq$  induced by  $Z$ . Whence, our joint total preorder  $\leq$  is not uniquely determined by  $\leq_i$ . Roughly speaking, the  $Z_i$  specify how the marginal preorders  $\leq_i$  scale relative to one another. As we shall see, this effectively means that our choice of  $Z_i$  affects the precision of our inferences: a good choice will ensure that any event of interest can be well approximated by elements of  $\Omega / \simeq$ . Of course, nothing prevents us, at least in theory, to consider the set of all  $Z_i$  which induce some given marginal total preorders  $\leq_i$ , and whence to work with a set of p-boxes. In some cases, this may result in quite complicated calculations. However, in Section 7.4, we will see an example where this approach is feasible.

Anyway, with this choice of  $Z$ , we can easily find the p-box which represents the joint as accurately as possible, under any rule of combination of coherent lower previsions:

**Lemma 22.** Consider any rule of combination  $\odot$  of coherent lower and upper previsions, mapping the marginals  $\underline{P}_1, \dots, \underline{P}_n$  to a joint coherent lower prevision  $\odot_{i=1}^n \underline{P}_i$  on all gambles. Suppose there are functions  $\ell$  and  $u$  for which:

$$\begin{aligned} \odot_{i=1}^n \underline{P}_i \left( \prod_{i=1}^n A_i \right) &= \ell(\underline{P}_1(A_1), \dots, \underline{P}_n(A_n)) \text{ and} \\ \odot_{i=1}^n \overline{P}_i \left( \prod_{i=1}^n A_i \right) &= u(\overline{P}_1(A_1), \dots, \overline{P}_n(A_n)), \end{aligned}$$

for all  $A_1 \subseteq \mathcal{X}_1, \dots, A_n \subseteq \mathcal{X}_n$ . Then, the couple  $(\underline{F}, \overline{F})$  defined by

$$\underline{F}(z) = \ell(\underline{F}_1(z), \dots, \underline{F}_n(z)) \quad \overline{F}(z) = u(\overline{F}_1(z), \dots, \overline{F}_n(z))$$

is the least conservative  $p$ -box on  $(\Omega, \leq)$  whose natural extension  $\underline{E}_{\underline{F}, \overline{F}}$  is dominated by the combination  $\odot_{i=1}^n \underline{P}_i$  of  $\underline{P}_1, \dots, \underline{P}_n$ .

*Proof.* By Theorem 2, the least conservative  $p$ -box on  $(\Omega, \leq)$  whose natural extension is dominated by the joint  $\underline{P} = \odot_{i=1}^n \underline{P}_i$  is given by

$$\underline{E}(z) = \underline{P}(Z^{-1}([0, z])) \quad \overline{F}(z) = \overline{P}(Z^{-1}([0, z]))$$

Now, observe that the set  $Z^{-1}([0, z])$  is a product of marginal intervals:

$$\begin{aligned} Z^{-1}([0, z]) &= \{(x_1, \dots, x_n) \in \Omega : \max_{i=1}^n Z_i(x_i) \leq z\} \\ &= \{(x_1, \dots, x_n) \in \Omega : (\forall i = 1, \dots, n)(Z_i(x_i) \leq z)\} \\ &= \prod_{i=1}^n \{x_i \in X_i : Z_i(x_i) \leq z\} = \prod_{i=1}^n Z_i^{-1}([0, z]). \end{aligned}$$

The desired equalities follow immediately.  $\square$

**7.2. Natural Extension: The Fréchet Case.** The natural extension  $\boxtimes_{i=1}^n \underline{P}_i$  of  $\underline{P}_1, \dots, \underline{P}_n$  is the lower envelope of all joint distributions (or, linear previsions) whose marginal distributions (or, marginal linear previsions) are compatible with the given marginal lower previsions. So, the model is completely vacuous about the dependence structure, as it includes all possible forms of dependence. We refer to for instance [8, p. 120, §3.1] for a rigorous definition. In this paper, we only need the following equalities, which are known as the Fréchet bounds [27] (see for instance [46, p. 131] for a more recent discussion):

$$\boxtimes_{i=1}^n \underline{P}_i \left( \prod_{i=1}^n A_i \right) = \max \left\{ 0, 1 - n + \sum_{i=1}^n \underline{P}_i(A_i) \right\} \text{ and} \quad (15a)$$

$$\boxtimes_{i=1}^n \overline{P}_i \left( \prod_{i=1}^n A_i \right) = \min_{i=1}^n \overline{P}_i(A_i) \quad (15b)$$

for all  $A_1 \subseteq \mathcal{X}_1, \dots, A_n \subseteq \mathcal{X}_n$ .

**Theorem 23.** The  $p$ -box  $(\underline{F}, \overline{F})$  defined by

$$\underline{F}(z) = \max \left\{ 0, 1 - n + \sum_{i=1}^n \underline{F}_i(z) \right\} \quad \overline{F}(z) = \min_{i=1}^n \overline{F}_i(z)$$

is the least conservative  $p$ -box on  $(\Omega, \leq)$  whose natural extension  $\underline{E}_{\underline{F}, \overline{F}}$  is dominated by the natural extension  $\boxtimes_{i=1}^n \underline{P}_i$  of  $\underline{P}_1, \dots, \underline{P}_n$ .

*Proof.* Immediate, by Lemma 22 and Eqs. (15).  $\square$

The next example shows that, even when  $\underline{P}_i$  are p-boxes, the joint p-box will in general only be an outer approximation (although the closest one that is a p-box) of the joint lower prevision.

*Example 24.* Consider two variables  $X$  and  $Y$  with domain  $\mathcal{X} = \{x_1, x_2\}$ , with  $x_1 < x_2$ , and  $\mathcal{Y} = \{y_1, y_2\}$ , with  $y_1 < y_2$ . Consider

$$\begin{aligned} \underline{F}_1(x_1) &= 0.4, & \overline{F}_1(x_1) &= 0.6, & \underline{F}_1(x_2) &= \overline{F}_1(x_2) = 1, \\ \underline{F}_2(y_1) &= 0.2, & \overline{F}_2(y_1) &= 0.3, & \underline{F}_2(y_2) &= \overline{F}_2(y_2) = 1. \end{aligned}$$

Let  $\underline{P}_1$  be the natural extension of  $(\underline{F}_1, \overline{F}_1)$ , and let  $\underline{P}_2$  be the natural extension of  $(\underline{F}_2, \overline{F}_2)$ . Consider the events  $A = \{x_1\} \times \mathcal{Y}$  and  $B = \mathcal{X} \times \{y_2\}$ . Writing  $\underline{P}$  for  $\boxtimes_{i=1}^n \underline{P}_i$ , we have that

$$\begin{aligned} \underline{P}(A) &= \underline{P}_1(\{x_1\}) = \max\{0, \underline{F}_1(x_1) - \overline{F}_1(x_1)\} = 0.4 \\ \underline{P}(B) &= \underline{P}_2(\{y_2\}) = \max\{0, \underline{F}_2(y_2) - \overline{F}_2(y_2)\} = 0.7 \end{aligned}$$

whence,

$$\begin{aligned} \underline{P}(A \cup B) &= \max\{\underline{P}(A), \underline{P}(B)\} = 0.7, \\ \underline{P}(A \cap B) &= \max\{0, 1 - 2 + \underline{P}(A) + \underline{P}(B)\} = 0.1. \end{aligned}$$

But this means that  $\underline{P}$  is not even 2-monotone, because  $\underline{P}(A \cup B) + \underline{P}(A \cap B) < \underline{P}(A) + \underline{P}(B)$ . Therefore,  $\underline{P}$  cannot be represented by a p-box, as p-boxes are completely monotone by Theorem 17.

**7.3. Independent Natural Extension.** In contrast, the *independent natural extension*  $\boxtimes_{i=1}^n \underline{P}_i$  of  $\underline{P}_1, \dots, \underline{P}_n$  models epistemic independence between  $X_1, \dots, X_n$ . We refer to [7] for a rigorous definition and properties. In this paper, we only need the following equalities:

$$\bigotimes_{i=1}^n \underline{P}_i \left( \prod_{i=1}^n A_i \right) = \prod_{i=1}^n \underline{P}_i(A_i) \text{ and} \tag{16a}$$

$$\bigotimes_{i=1}^n \overline{P}_i \left( \prod_{i=1}^n A_i \right) = \prod_{i=1}^n \overline{P}_i(A_i) \tag{16b}$$

for all  $A_1 \subseteq \mathcal{X}_1, \dots, A_n \subseteq \mathcal{X}_n$ .

**Theorem 25.** *The p-box  $(\underline{F}, \overline{F})$  defined by*

$$\underline{F}(z) = \prod_{i=1}^n \underline{F}_i(z) \qquad \overline{F}(z) = \prod_{i=1}^n \overline{F}_i(z)$$

*is the least conservative p-box on  $(\Omega, \leq)$  whose natural extension  $\underline{E}_{\underline{F}, \overline{F}}$  is dominated by the independent natural extension  $\boxtimes_{i=1}^n \underline{P}_i$  of  $\underline{P}_1, \dots, \underline{P}_n$ .*

*Proof.* Immediate, by Lemma 22 and Eqs. (16).  $\square$

Again, in general, the joint p-box will only be an outer approximation of the actual joint lower prevision.

*Example 26.* Again, consider two variables  $X$  and  $Y$  with domain  $\mathcal{X} = \{x_1, x_2\}$ , with  $x_1 < x_2$ , and  $\mathcal{Y} = \{y_1, y_2\}$ , with  $y_1 < y_2$ . Consider

$$\begin{aligned} \underline{F}_1(x_1) &= 0.4, & \overline{F}_1(x_1) &= 0.6, & \underline{F}_1(x_2) &= \overline{F}_1(x_2) = 1, \\ \underline{F}_2(y_1) &= 0.3, & \overline{F}_2(y_1) &= 0.5, & \underline{F}_2(y_2) &= \overline{F}_2(y_2) = 1. \end{aligned}$$

As before, let  $\underline{P}_1$  be the natural extension of  $(\underline{E}_1, \overline{F}_1)$ , and let  $\underline{P}_2$  be the natural extension of  $(\underline{E}_2, \overline{F}_2)$ . Consider the events  $A = \{(x_1, y_1), (x_1, y_2)\}$  and  $B = \{(x_1, y_2), (x_2, y_1)\}$ . Writing  $\underline{P}$  for  $\otimes_{i=1}^n \underline{P}_i$ , we have that

$$\begin{aligned}\underline{P}(A) &= \underline{P}_1(\{x_1\}) = 0.4 \\ \underline{P}(B) &\geq 0.4\end{aligned}$$

where the last inequality follows from the fact that all probability mass functions  $p$  which dominate  $\underline{P}$  must satisfy  $p(x_1|y_2) \geq \underline{P}(\{x_1\}) = 0.4$  and  $p(x_2|y_1) \geq \underline{P}(\{x_2\}) = 0.4$ , whence

$$p(B) = p(x_1|y_2)p(y_2) + p(x_2|y_1)p(y_1) \geq 0.4(p(y_1) + p(y_2)) = 0.4$$

for all  $p$  which dominate  $\underline{P}$ . Because  $\underline{P}$  is the lower envelope of all such  $p$ , the desired inequality follows.<sup>8</sup> Also, because of the factorization property of the independent natural extension,

$$\begin{aligned}\underline{P}(A \cup B) &= 1 - \overline{P}(\{(x_2, y_2)\}) = 1 - \overline{P}_1(\{x_2\})\overline{P}_2(\{y_2\}) = 1 - 0.6 \times 0.7 = 0.58 \\ \underline{P}(A \cap B) &= \underline{P}(\{(x_1, y_2)\}) = \underline{P}_1(\{x_1\})\underline{P}_2(\{y_2\}) = 0.4 \times 0.5 = 0.2.\end{aligned}$$

Again, this means that  $\underline{P}$  cannot be represented by a p-box, as it violates 2-monotonicity.

**7.4. Special Case: Probabilistic Arithmetic.** Let  $Y = X_1 + X_2$  with  $X_1$  and  $X_2$  real-valued random variables. One can also consider subtraction, multiplication, and division, but for simplicity, we stick to addition—the other three cases follow along almost identical lines.

Probabilistic arithmetic [46] deals with the problem of estimating  $\underline{P}_Y([-\infty, y]) = \underline{F}_Y(y)$  and  $\overline{P}_Y([-\infty, y]) = \overline{F}_Y(y)$  for any  $y \in \mathbb{R}$  under the assumptions that

- the uncertainty on  $X_1$  and  $X_2$  is given by p-boxes  $(\underline{E}_1, \overline{F}_1)$  and  $(\underline{E}_2, \overline{F}_2)$ , with  $\leq_1$  and  $\leq_2$  the natural ordering of real numbers,<sup>9</sup> and
- the dependence structure is completely unknown (Fréchet case).

Using the Fréchet bounds, Williamson and Downs [45] provide explicit formulae for the different arithmetic operations, thus providing very efficient algorithms to make inferences from marginal p-boxes.

Let us show, for the particular case of addition, that their results are captured by our joint p-box proposed in Theorem 23. Other cases, not treated here to save space, follow from almost identical reasoning.<sup>10</sup> The lower cumulative distribution function resulting from probabilistic arithmetic is, for any  $y \in \mathbb{R}$

$$\underline{F}_{X_1+X_2}(y) = \sup_{x_1, x_2: x_1+x_2=y} \max\{0, \underline{E}_1(x_1) + \underline{E}_2(x_2) - 1\}. \quad (17)$$

Without much loss of generality, and for our convenience, assume that both  $X_1$  and  $X_2$  lie in a bounded interval  $[a, b]$ .

Let  $Z_1$  and  $Z_2$  be any surjective maps  $[a, b] \rightarrow [0, 1]$  which induce the usual ordering on  $[0, 1]$ . Some properties of  $Z_1$  and  $Z_2$  immediately follow: both are continuous and strictly increasing, and so are their inverses—we rely on this in a bit.

To apply Theorem 23, we consider the total preorder  $\leq$  on  $\Omega = [a, b]^2$  induced by  $Z(x_1, x_2) = \max\{Z_1(x_1), Z_2(x_2)\}$ . Figure 3 illustrates the event<sup>11</sup>  $\{X_1 + X_2 \leq y\}$ , with  $y \in [2a, 2b]$ , as well as the largest interval  $Z^{-1}([0, z])$  included in it. Recall that

$$Z^{-1}([0, z]) = Z_1^{-1}([0, z]) \times Z_2^{-1}([0, z]) = [0, Z_1^{-1}(z)] \times [0, Z_2^{-1}(z)].$$

<sup>8</sup>By linear programming, it can actually be shown that  $\underline{P}(B) = 0.4$ .

<sup>9</sup>For subtraction and division,  $\leq_2$  is the reverse natural ordering.

<sup>10</sup>Note that  $X_1$  and  $X_2$  are assumed to be positive in case of multiplication and division.

<sup>11</sup>By  $\{X_1 + X_2 \leq y\}$  we mean  $\{(x_1, x_2) \in [0, 1]^2: x_1 + x_2 \leq y\}$ .

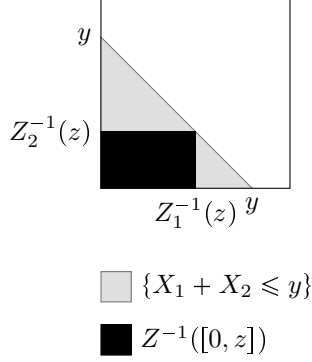


FIGURE 3. The event  $\{X_1 + X_2 \leq y\}$ , and the largest interval  $Z^{-1}([0, z])$  included in it.

Whence, for  $z$  such that  $Z_1^{-1}(z) + Z_2^{-1}(z) = y$ , we achieve the largest interval  $Z^{-1}([0, z])$  which is still included in  $\{X_1 + X_2 \leq y\}$ . There is always a unique such  $z$  because also  $Z_1^{-1} + Z_2^{-1}$  is continuous and strictly increasing.

Recall that, by Theorem 23 (now without shortcuts in notation),

$$\begin{aligned} \underline{F}(Z^{-1}(z)) &= \max\{0, \underline{F}_1(Z_1^{-1}(z)) + \underline{F}_2(Z_2^{-1}(z)) - 1\} \\ \overline{F}(Z^{-1}(z)) &= \min_{i=1}^n \{\overline{F}_1(Z_1^{-1}(z)), \overline{F}_2(Z_2^{-1}(z))\} \end{aligned}$$

is the least conservative p-box on  $(\Omega, \leq)$  whose natural extension is dominated by the natural extension  $\underline{P}_1 \boxtimes \underline{P}_2$  of  $\underline{P}_1$  and  $\underline{P}_2$ .

Also, as we have just shown,  $Z^{-1}([0, z])$ , for our choice of  $z$ , is the topological interior of  $\{X_1 + X_2 \leq y\}$ . Whence, by Theorem 20, we find that

$$\begin{aligned} \underline{E}_{\underline{F}, \overline{F}}(\{X_1 + X_2 \leq y\}) &= \underline{E}_{\underline{F}, \overline{F}}(Z^{-1}([0, z])) = \underline{F}(Z^{-1}(z)) \\ &= \max\{0, \underline{F}_1(Z_1^{-1}(z)) + \underline{F}_2(Z_2^{-1}(z)) - 1\} \end{aligned}$$

where we remember that  $z$  is chosen such that  $Z_1^{-1}(z) + Z_2^{-1}(z) = y$ .

But, this holds for every valid choice of functions  $Z_1$  and  $Z_2$ , whence

$$\underline{P}_1 \boxtimes \underline{P}_2(\{X_1 + X_2 \leq y\}) \geq \sup_{x_1, x_2: x_1 + x_2 = y} \max\{0, \underline{F}_1(x_1) + \underline{F}_2(x_2) - 1\}$$

which indeed coincides with Eq. (17). Similar arguments hold for the upper cumulative distribution functions, and other arithmetic operations.

In conclusion, probabilistic arithmetic constitutes a very specific case of our approach.

## 8. EXAMPLES

In this section, we investigate two different examples in which p-boxes are used to model uncertainty around some parameters. The first example concerns a damped harmonic oscillator, i.e., a classical engineering toy example. The second example concerns the evaluation of a river dike height, an important issue in regions subject to potential floods.

**8.1. Damped Harmonic Oscillator.** Consider a simple damped harmonic oscillator, with damping coefficient  $c > 0$ , spring constant  $k > 0$ , and mass  $m > 0$ . The *damping ratio*

$$\zeta(c, k) = \frac{c}{2\sqrt{km}}$$

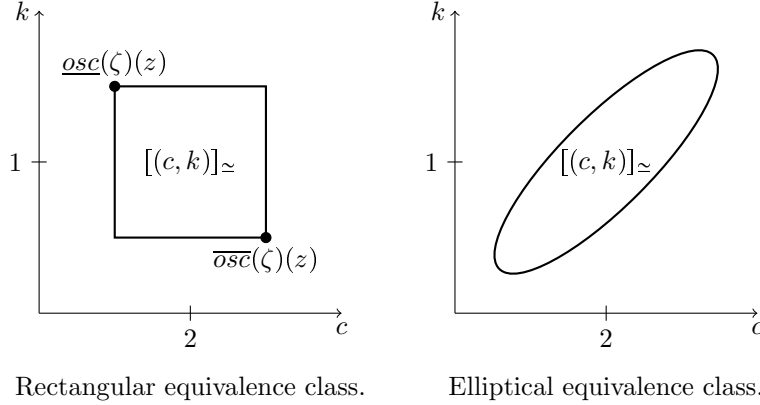


FIGURE 4. Different possible equivalence classes.

determines how quickly the oscillator returns to its equilibrium state— $\zeta(c, k) = 1$  means fastest convergence. Suppose the engineering design has already been completed, so the optimal values for  $c^*$  and  $k^*$  have been determined, such that  $\zeta(c^*, k^*) = 1$ .

Without loss of generality, we choose the units for mass, time, and length, such that  $m = k^* = \zeta(c^*, k^*) = 1$  (so  $c^* = 2$ ).

Of course, the actual values for  $c$  and  $k$  will differ from their design values, and uncertainty must be taken into account. Let us calculate the lower and upper expectation of  $\zeta(c, k)$ , given that our uncertainty about  $c$  and  $k$  is described by a p-box.

First, we must specify a preorder. For this problem, it seems fairly natural to have bounds on the quantiles of the distance between the actual values  $(c, k)$  and the design values  $(2, 1)$ . This comes down to for instance the following choice for  $Z$ :

$$Z(c, k) = \max\{|c - 2|, 2|k - 1|\}$$

For simplicity, we only consider the region  $Z(c, k) \leq 1$ . This means that we are certain that  $c \in [1, 3]$  and  $k \in [0.5, 1.5]$ —if necessary,  $Z$  can be rescaled to accommodate larger or smaller regions. Note that we have taken a supremum norm as distance. This simplifies the calculations below, but of course, one might as well take the Euclidian norm, or any other reasonable distance function, at the expense of slightly more complicated calculations and dependency modelling (see Fig. 4).

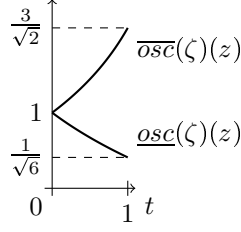
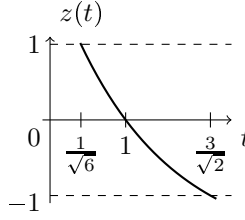
Equivalence classes  $[(c, k)]_{\approx}$  are edges of rectangles with vertices

$$(2 \pm Z(c, k), 1 \pm Z(c, k)/2).$$

What is a p-box for the preorder  $\leq$  induced by  $Z$ ? A p-box  $(\underline{F}, \overline{F})$  specifies lower and upper bounds for the probability of concentric rectangles around the design point  $(2, 1)$ :

$$\underline{F}(z) \leq p(\{(c, k) : Z(c, k) \leq z\}) \leq \overline{F}(z)$$

So, effectively, our p-box specifies concentric prediction regions for the uncertain parameters  $c$  and  $k$ .

FIGURE 5. The lower oscillation  $\underline{osc}(\zeta)(z)$  and upper oscillation  $\overline{osc}(\zeta)(z)$ .FIGURE 6. The function  $z(t)$  which determines the cut sets.

We can now calculate the lower and upper expectation of  $\zeta(c, k)$ . First, we calculate the lower oscillation  $\underline{osc}(\zeta)$  and upper oscillation  $\overline{osc}(\zeta)$  (see Fig. 5):

$$\begin{aligned}\underline{osc}(\zeta)(z) &= \inf_{(c,k): Z(c,k)=z} \zeta(c,k) = \frac{2-z}{2\sqrt{(1+z/2)}} \\ \overline{osc}(\zeta)(z) &= \sup_{(c,k): Z(c,k)=z} \zeta(c,k) = \frac{2+z}{2\sqrt{(1-z/2)}}\end{aligned}$$

Next, we find the full components of the events

$$\begin{aligned}L_t &= \{z \in [0, 1]: \underline{osc}(\zeta)(z) \geq t\} = \left\{z \in [0, 1]: \frac{2-z}{2\sqrt{(1+z/2)}} \geq t\right\} \\ U_t &= \{z \in [0, 1]: \overline{osc}(\zeta)(z) \geq t\} = \left\{z \in [0, 1]: \frac{2+z}{2\sqrt{(1-z/2)}} \geq t\right\}\end{aligned}$$

for all  $t \in [0, 0.5]$ . Fortunately,  $\underline{osc}(\zeta)$  is decreasing as function of  $z$ , and  $\overline{osc}(\zeta)$  is increasing, and hence  $L_t = [0, \ell_t]$  and  $U_t = [u_t, 1]$ , with (see Fig. 6)

$$\begin{aligned}\ell_t &= 2 - t(-t + \sqrt{t^2 + 8}) := z(t) && \text{with } t \in [\frac{1}{\sqrt{6}}, 1] \\ u_t &= -2 + t(-t + \sqrt{t^2 + 8}) := -z(t) && \text{with } t \in [1, \frac{3}{\sqrt{2}}]\end{aligned}$$

Note that the given bounds for  $t$  arise from the minimum and maximum of  $\underline{osc}(f)$  and  $\overline{osc}(f)$ . Concluding, by Proposition 21, when  $\underline{E}(z) = \underline{E}(z-)$  for all  $z \in [0, 1]$  and  $\underline{E}(0) = 0$ ,

$$\begin{aligned}\underline{E}(\zeta) &= \frac{1}{\sqrt{6}} + \int_{\frac{1}{\sqrt{6}}}^1 \underline{E}(z(t)) dt \\ \overline{E}(\zeta) &= 1 + \int_1^{3/\sqrt{2}} (1 - \underline{E}(-z(t))) dt\end{aligned}$$



symbol	name	unit
$h$	overflow height of the river	$m$
$q$	river flow rate	$m^3 s^{-1}$
$b$	river width	$m$
$k$	Strickler coefficient	$m^{1/3} s^{-1}$
$u$	upriver water level	$m$
$d$	downriver water level	$m$
$\ell$	length of river stretch	$m$

TABLE 1. Meaning of the variables used in Eq. (18)

Interestingly, both the lower and upper expectation of  $\zeta$  are determined by the lower cumulative distribution function only. Hence, in this problem, we actually do not need to elicit the upper cumulative distribution function.<sup>12</sup>

For example, if the expert says that  $c$  and  $k$  are independent, and the marginal lower cumulative distribution functions are uniform on  $c \in [1, 2]$  and  $k \in [0.5, 1.5]$ , so  $\underline{F}_1(z) = \underline{F}_2(z) = z$ , with preorders induced by  $Z_1(c) = |c - 2|$  and  $Z_2(k) = 2|k - 1|$ , then, by Theorem 25, because  $Z = \max\{Z_1, Z_2\}$ , it follows that  $\underline{F}(z) = z^2$ , and

$$\begin{aligned}\underline{E}(\zeta) &= \frac{1}{\sqrt{6}} + \int_{\frac{1}{\sqrt{6}}}^1 z(t)^2 dt = 0.584 \\ \overline{E}(\zeta) &= 1 + \int_1^{3/\sqrt{2}} (1 - z(t)^2) dt = 1.664\end{aligned}$$

**8.2. River Dike Height Estimation.** We aim to estimate the minimal required dike height along a given stretch of river, using a simplified model that is used by the EDF (the French integrated energy operator) to make initial evaluations [13]. Although this model is quite simple, it provides a realistic industrial application. Skipping technical details, the model results in the following relationship:

$$h(q, k, u, d) = \begin{cases} \left( \frac{q}{k \sqrt{\frac{u-d}{\ell} b}} \right)^{\frac{3}{5}} & \text{if } q \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (18)$$

The meaning of the variables is summarised in Table 1.

For the particular case under study, the river width is  $b = 300m$  and the river length is  $\ell = 6400m$ . The remaining parameters are uncertain. Expert assessment leads to the following distributions:

- The maximal flow rate  $q$  has a Gumbel distribution<sup>13</sup> with location parameter  $\mu = 1335m^3 s^{-1}$  and scale parameter  $\beta = 716m^3 s^{-1}$ . For calculations, it is easier to work with symmetric distributions. Therefore, we introduce a variable  $r$  satisfying

$$q = \mu - \beta \ln(-\ln(r)).$$

<sup>12</sup>Of course this will not always be the case—it just happens to be so for this example.

<sup>13</sup>The Gumbel distribution models the maximum of an exponentially distributed sample, and is used in extreme value theory [12] to model rare events such as floods.

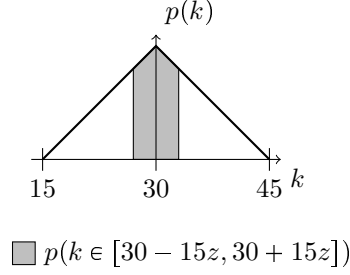


FIGURE 7. Derivation of the p-box for a triangular distribution.

If  $r$  is uniformly distributed over  $[0, 1]$ , then  $q$  has the Gumbel distribution with location parameter  $\mu$  and scale parameter  $\beta$ . So, after transformation,

$$h(r, k, u, d) = \begin{cases} \left( \frac{\mu - \beta \ln(-\ln(r))}{k \sqrt{\frac{u-d}{\epsilon} b}} \right)^{\frac{3}{5}} & \text{if } \mu - \beta \ln(-\ln(r)) \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

- The Strickler coefficient  $k$  has a symmetric triangular distribution over the interval  $[15m^{1/3}s^{-1}, 45m^{1/3}s^{-1}]$  (with mode at  $k^* = 30m^{1/3}s^{-1}$ ).
- There is also uncertainty about the water levels  $u$  and  $d$ , because sedimentary conditions are hard to characterise. Measured values are  $u^* = 55m$  and  $d^* = 50m$ , with measurement error definitely less than  $1m$ . These are also modelled by symmetric triangular distributions, on  $[54m, 56m]$  and  $[49m, 51m]$  respectively.

Again, a natural choice for  $Z$  is the distance between the expected values ( $r^* = 1/2, k^* = 30, u^* = 55, d^* = 50$ ) and the actual values  $(r, k, u, d)$ :

$$Z(r, k, u, d) = \max\{2|r - 1/2|, |k - 30|/15, |u - 55|, |d - 50|\},$$

The scale of the distances has been chosen such that  $Z(r, k, u, d) \leq 1$  for all points of interest. Equivalence classes  $[(r, k, u, d)]_{\simeq}$  are borders of 4-dimensional boxes with vertices

$$((1 \pm z)/2, 30 \pm 15z, 55 \pm z, 50 \pm z)$$

where  $z = Z(r, k, u, d)$ .

The marginal p-boxes are, for  $r$ :

$$\underline{F}_1(z) = \overline{F}_1(z) = p(2|r - 1/2| \leq z) = p(r \in [(1 - z)/2, (1 + z)/2]) = z$$

because  $r$  is uniformly distributed over  $[0, 1]$ . For  $k$ , we have:

$$\underline{F}_2(z) = \overline{F}_2(z) = p(|k - 30|/15 \leq z) = p(k \in [30 - 15z, 30 + 15z]) = 1 - (1 - z)^2$$

(see Fig. 7). Similarly, for  $u$  and  $d$ , it is easily verified that:

$$\underline{F}_3(z) = \overline{F}_3(z) = \underline{F}_4(z) = \overline{F}_4(z) = 1 - (1 - z)^2$$

The lower oscillation  $\underline{osc}(h)$  and upper oscillation  $\overline{osc}(h)$  can be calculated along the same lines as in the previous example:

$$\begin{aligned} \underline{osc}(h)(z) &= \inf_{(r, k, u, d) : Z(r, k, u, d) = z} h(r, k, u, d) = o(-z) \\ \overline{osc}(h)(z) &= \sup_{(r, k, u, d) : Z(r, k, u, d) = z} h(r, k, u, d) = o(z) \end{aligned}$$

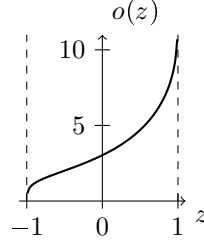


FIGURE 8. The function  $o(z)$  which determines the lower and upper oscillation, and the cut sets.

with

$$o(z) = \begin{cases} \left( \frac{\mu - \beta \ln(-\ln((1+z)/2))}{(30-15z)\sqrt{\frac{5-2z}{t}b}} \right)^{\frac{3}{5}} & \text{if } \mu - \beta \ln(-\ln((1+z)/2)) \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The function  $o(z)$  is depicted in Fig. 8: it is increasing, with  $o(-1) = 0$  (this is not immediately clear from the picture, but at higher scale, it becomes apparent),  $o(0) = 3.032$ , and  $o(1) = +\infty$ .

Again,  $\underline{osc}(h)(z)$  and  $\overline{osc}(h)(z)$  are decreasing and increasing in  $z$ , respectively. Hence the full components of the events

$$\begin{aligned} L_t &= \{z \in [0, 1] : \underline{osc}(h)(z) \geq t\} = \{z \in [0, 1] : o(-z) \geq t\} \\ U_t &= \{z \in [0, 1] : \overline{osc}(h)(z) \geq t\} = \{z \in [0, 1] : o(z) \geq t\} \end{aligned}$$

are of the form  $L_t = [0, \ell_t]$  and  $U_t = [u_t, 1]$  again, with

$$\ell_t = -o^{-1}(t) \text{ for } t \leq o(0) \quad \quad u_t = o^{-1}(t) \text{ for } t \geq o(0)$$

As in the previous example, we do not need to elicit the upper cumulative distributions, and only the lower ones need to be given. With unknown dependence, using Theorem 23, we have

$$\underline{F}(z) = \max\{0, -3 + z + 3(1 - (1 - z)^2)\}$$

and whence

$$\begin{aligned} \underline{E}(h) &= \int_0^{o(0)} \underline{F}(-o^{-1}(t)) dt = 1.515 \\ \overline{E}(h) &= o(0) + \int_{o(0)}^{+\infty} (1 - \underline{F}(o^{-1}(t))) dt = 6.423 \end{aligned}$$

Therefore, to be on the safe side, we should consider average overflowing heights of at least  $6.5m$ . For comparison, using traditional methods instead of p-boxes,  $h$  has expectation  $3.2m$ , assuming independence between all variables—this lies between the lower and upper expectation that we just calculated, as expected. The interval is obviously much wider:

- because we have reduced a multivariate problem to a univariate one, whence, leading to imprecision due to the difference between lower and upper oscillation,
- and because we have not made any assumption of independence, whence, leading to imprecision due to weaker assumptions.

Realistically, the decision maker may desire a dike height  $t$  such that the upper probability of disaster  $\overline{E}(\{h \geq t\})$  is less than a given threshold. It is easily verified that:

$$\overline{E}(\{h \geq t\}) = \overline{E}(\{\overline{osc}(h) \geq t\}) = 1 - \underline{F}(o^{-1}(t))$$

For instance, for  $\overline{E}(\{h \geq t\}) = 0.01$ , we need  $t$  to be  $10.725m$ . For comparison, using traditional methods instead of p-boxes,  $t$  needs to be about  $9m$ , assuming independence between all variables.

In both examples, analytical calculations are relatively simple due to the monotonicity of the target function with respect to the uncertain variables. Of course, this may not be the case in general.

## 9. CONCLUSIONS

P-boxes are one of the most interesting imprecise probability models from an operational point of view, because they are simply characterised by a lower and an upper cumulative distribution function. In this paper, for the purpose of multivariate modelling, we studied inferences (lower and upper expectations in particular) from p-boxes on arbitrary totally preordered spaces. For this purpose, we represented p-boxes as coherent lower previsions, and studied their natural extension.

We used an as general as possible model by considering p-boxes whose lower and upper cumulative distribution functions are defined on a totally preordered space. Thereby, we extended the theory of p-boxes from finite to infinite sets, and from total orders to total preorders. This allowed us to unify p-boxes on finite spaces and on intervals of real numbers, and to extend the theory to the multivariate case.

One very interesting result of this paper is a practical means of calculating the natural extension of a p-box in this general setting. We proved that the natural extension of a p-box is arbitrarily additive on full components of clopen sets with respect to the partition topology induced by equivalence classes of the underlying preorder (Theorem 13, Corollaries 15 and 16). We also proved that the natural extension is completely monotone, and therefore has a Choquet integral representation (Theorem 18). Consequently, to calculate the natural extension, we proved that it suffices to calculate the full components of the cut sets of the lower oscillation, followed by a simple Riemann integral (Proposition 19).

As a special case, we studied p-boxes whose preorders are induced by a real-valued mapping. Such p-boxes are particularly attractive, as they allow to build or elicit a multivariate uncertainty model at once. They correspond to lower and upper probabilistic bounds given over nested regions that can take arbitrary shapes.

Consequently, we provided a new tool to combine marginal p-boxes into a joint p-box, under arbitrary rules of combination, thereby allowing any type of dependency modelling (Lemma 22). As examples, we considered two extreme cases: assuming nothing about dependence (Theorem 23), and assuming epistemic independence (Theorem 25). Similar formulas are easily derived for any other rule of combination. Moreover, Williamson and Downs's [45] probabilistic arithmetic obtains as a special case of our approach.

We demonstrated our methodology on inference about a damped harmonic oscillator, and on a river dike assessment, showing that calculations are generally straightforward.

Of course, many open problems regarding p-boxes remain to be investigated. For instance, even though there need not be any relation between the preorder and the dependency model—because one can, in theory at least, always construct a multivariate p-box from marginals for any dependency model and any preorder—some combinations obviously lead to more imprecision than others. Our choice led to simple mathematical expressions, but is perhaps not the best one possible in terms of precision. Can the dependency model inform the choice of preorder, to arrive at tighter bounds? Also, the connection of p-boxes with other uncertainty models, such as possibility measures and clouds, deserves further investigation.

## ACKNOWLEDGEMENTS

Work partially supported by a doctoral grant from the IRSN. We are particularly grateful to Enrique Miranda for the many very fruitful discussions, extremely useful suggestions, and various contributions to this paper. We also thank Gert de Cooman and Didier Dubois for their help with a very early draft of this paper. Finally, we thank both reviewers, whose constructive comments helped improving the presentation of the paper substantially.

## REFERENCES

- [1] Ph. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.
- [2] Y. Ben-Haim. *Info-gap decision theory: decisions under severe uncertainty*. Academic Press, London, 2006.
- [3] C. Berge. *Espaces topologiques, fonctions multivoques*. Dunod, Paris, 1959.
- [4] G. Boole. *An investigation of the laws of thought on which are founded the mathematical theories of logic and probabilities*. Walton and Maberly, London, 1854.
- [5] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953–1954.
- [6] G. de Cooman. Possibility theory I–III. *International Journal of General Systems*, 25:291–371, 1997.
- [7] G. de Cooman, E. Miranda, and M. Zaffalon. Independent natural extension. In Eyke Hüllermeier, Rudolf Kruse, and Frank Hoffmann, editors, *Computational Intelligence for Knowledge-Based Systems Design*, Lecture Notes in Computer Science, pages 737–746. Springer, 2010.
- [8] G. de Cooman and M. C. M. Troffaes. Coherent lower previsions in systems modelling: products and aggregation rules. *Reliability Engineering and System Safety*, 85:113–134, 2004.
- [9] G. de Cooman, M. C. M. Troffaes, and E. Miranda.  $n$ -Monotone exact functionals. *Journal of Mathematical Analysis and Applications*, 347(1):143–156, 2008.
- [10] G. de Cooman, M. C. M. Troffaes, and E. Miranda. A unifying approach to integration for bounded positive charges. *Journal of Mathematical Analysis and Applications*, 340(2):982–999, 2008.
- [11] B. de Finetti. *Theory of Probability*. John Wiley & Sons, Chichester, 1974–1975.
- [12] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer, 2006.
- [13] E. de Rocquigny, N. Devictor, and S. Tarantola. *Uncertainty in Industrial Practice: A guide to Quantitative Uncertainty Management*, chapter 10. Partial Safety Factors to Deal with Uncertainties in Slope Stability of River Dykes. Wiley, 2008.
- [14] G. Debreu. Integration of correspondences. In *Fifth Berkeley Symposium of Mathematical Statistics and Probability*, pages 351–372, Berkeley (USA), 1965.
- [15] F. Delbaen. Coherent risk measures on general probability spaces. In K. Sandmann and P. J. Schönbucher, editors, *Advances in Finance and Stochastics*, pages 1–37. Delbaen, F, Berlin, 2002.
- [16] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.
- [17] S. Destercke and D. Dubois. A unified view of some representations of imprecise probabilities. In J. Lawry, E. Miranda, A. Bugarín, S. Li, M.A. Gil, P. Gregorzewski, and O. Hryniewiczet, editors, *Soft Methods for Integrated Uncertainty Modelling*, Advances in Soft Computing, pages 249–257, Bristol, 2006. Springer.
- [18] S. Destercke and D. Dubois. The role of generalised p-boxes in imprecise probability models. In Thomas Augustin, Frank P. A. Coolen, Serafin Moral, and Matthias C. M. Troffaes, editors, *ISIPTA '09: Proceedings of the Sixth International Symposium on Imprecise Probability: Theories and Applications*, pages 179–188, Durham, UK, July 2009. SIPTA.
- [19] S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations: I. Generalized p-boxes. *International Journal of Approximate Reasoning*, 49(3):649–663, 2008.
- [20] S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations: II. Clouds. *International Journal of Approximate Reasoning*, 49(3):664–677, 2008.
- [21] S. Destercke and O. Strauss. Using cloudy kernels for imprecise linear filtering. In *IPMU*, pages 198–207, 2010.
- [22] D. Dubois and H. Prade. The mean value of a fuzzy number. *Fuzzy Sets and Systems*, 24(3):279–300, 1987.
- [23] D. Dubois and H. Prade. *Possibility Theory*. Plenum Press, New York, 1988.
- [24] S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002–4015, Sandia National Laboratories, January 2003.
- [25] S. Ferson and W. Tucker. Sensitivity analysis using probability bounding. *Reliability engineering and system safety*, 91(10-11):1435–1442, 2006.

- [26] S. Ferson and W. Tucker. Probability boxes as info-gap models. In *Proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society*, New York (USA), 2008.
- [27] M. Fréchet. Généralisations du théorème des probabilités totales. *Fundamenta Mathematica*, 25:379–387, 1935.
- [28] M. Fuchs and A. Neumaier. Potential based clouds in robust design optimization. *Journal of Statistical Theory and Practice, Special Issue on Imprecision*, 3(1):225–238, 2008.
- [29] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963.
- [30] E. Kriegler and H. Held. Utilizing belief functions for the estimation of future climate change. *International Journal of Approximate Reasoning*, 39:185–209, 2005.
- [31] I. Levi. *The Enterprise of Knowledge*. MIT Press, London, 1980.
- [32] E. Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48(2):628–658, 2008.
- [33] E. Miranda, G. de Cooman, and E. Quaeghebeur. Finitely additive extensions of distribution functions and moment sequences: the coherent lower prevision approach. *International Journal of Approximate Reasoning*, 48(1):132–155, 2008.
- [34] H. T. Nguyen. On random sets and belief functions. *Journal of Mathematical Analysis and Applications*, 65(3):531–542, 1978.
- [35] M. Oberguggenberger and W. Fellin. Reliability bounds through random sets: non-parametric methods and geotechnical applications. *Computers and Structures*, 86(10):1093–1101, 2008.
- [36] H.M. Regan, S. Ferson, and D. Berleant. Equivalence of methods for uncertainty propagation of real-valued random variables. *International Journals of Approximate Reasoning*, 36:1–30, 2004.
- [37] E. Schechter. *Handbook of Analysis and Its Foundations*. Academic Press, San Diego, CA, 1997.
- [38] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.
- [39] M. C. M. Troffaes. *Optimality, Uncertainty, and Dynamic Programming with Lower Previsions*. PhD thesis, Ghent University, Ghent, Belgium, March 2005.
- [40] E. Tsiporkova-Hristoskova, B. De Baets, and E. Kerre. A fuzzy inclusion based approach to upper inverse images under fuzzy multivalued mappings. *Fuzzy Sets and Systems*, 85(1):93–108, 1997.
- [41] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
- [42] K. Weichselberger. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I — Intervallwahrscheinlichkeit als umfassendes Konzept*. Physica, Heidelberg, 2001. In cooperation with T. Augustin and A. Wallner.
- [43] P. M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975.
- [44] P. M. Williams. Notes on conditional previsions. *International Journal of Approximate Reasoning*, 44:366–383, 2007. Revised journal version of [43].
- [45] R. C. Williamson and T. Downs. Probabilistic arithmetic I: Numerical methods for calculating convolutions and dependency bounds. *International Journal of Approximate Reasoning*, 4:8–158, 1990.
- [46] R.C. Williamson. *Probabilistic Arithmetic*. PhD thesis, University of Queensland, Australia, 1989.

DURHAM UNIVERSITY, DEPT. OF MATHEMATICAL SCIENCES, SCIENCE LABORATORIES, SOUTH ROAD, DURHAM DH1 3LE, UNITED KINGDOM

*E-mail address:* `matthias.troffaes@gmail.com`

CIRAD, UMR1208, 2 PLACE P. VIALA, F-34060 MONTPELLIER CEDEX 1, FRANCE

*E-mail address:* `sebastien.destercke@cirad.fr`